

(19) World Intellectual Property Organization  
International Bureau



(43) International Publication Date  
20 September 2001 (20.09.2001)

PCT

(10) International Publication Number  
**WO 01/68848 A2**

- (51) International Patent Classification<sup>7</sup>: **C12N 15/12**,  
15/62, C07K 14/47, 14/705, 16/18, G01N 33/53, C12Q  
1/68
- (21) International Application Number: PCT/US01/06520
- (22) International Filing Date: 28 February 2001 (28.02.2001)
- (25) Filing Language: English
- (26) Publication Language: English
- (30) Priority Data:
- |                |                             |    |
|----------------|-----------------------------|----|
| PCT/US00/05601 | 1 March 2000 (01.03.2000)   | US |
| PCT/US00/05841 | 2 March 2000 (02.03.2000)   | US |
| 60/187,202     | 3 March 2000 (03.03.2000)   | US |
| 60/186,968     | 6 March 2000 (06.03.2000)   | US |
| 60/189,328     | 14 March 2000 (14.03.2000)  | US |
| 60/189,320     | 14 March 2000 (14.03.2000)  | US |
| PCT/US00/06884 | 15 March 2000 (15.03.2000)  | US |
| 60/191,048     | 21 March 2000 (21.03.2000)  | US |
| 60/190,828     | 21 March 2000 (21.03.2000)  | US |
| 60/191,314     | 21 March 2000 (21.03.2000)  | US |
| 60/191,007     | 21 March 2000 (21.03.2000)  | US |
| 60/192,655     | 28 March 2000 (28.03.2000)  | US |
| 60/193,032     | 29 March 2000 (29.03.2000)  | US |
| 60/193,053     | 29 March 2000 (29.03.2000)  | US |
| PCT/US00/08439 | 30 March 2000 (30.03.2000)  | US |
| 60/194,647     | 4 April 2000 (04.04.2000)   | US |
| 60/194,449     | 4 April 2000 (04.04.2000)   | US |
| 60/196,820     | 11 April 2000 (11.04.2000)  | US |
| 60/195,975     | 11 April 2000 (11.04.2000)  | US |
| 60/196,000     | 11 April 2000 (11.04.2000)  | US |
| 60/196,187     | 11 April 2000 (11.04.2000)  | US |
| 60/196,690     | 11 April 2000 (11.04.2000)  | US |
| 60/198,121     | 18 April 2000 (18.04.2000)  | US |
| 60/198,585     | 18 April 2000 (18.04.2000)  | US |
| 60/199,654     | 25 April 2000 (25.04.2000)  | US |
| 60/199,397     | 25 April 2000 (25.04.2000)  | US |
| 60/199,550     | 25 April 2000 (25.04.2000)  | US |
| 60/201,516     | 3 May 2000 (03.05.2000)     | US |
| PCT/US00/13705 | 17 May 2000 (17.05.2000)    | US |
| PCT/US00/14042 | 22 May 2000 (22.05.2000)    | US |
| PCT/US00/14941 | 30 May 2000 (30.05.2000)    | US |
| PCT/US00/15264 | 2 June 2000 (02.06.2000)    | US |
| 60/209,832     | 5 June 2000 (05.06.2000)    | US |
| PCT/US00/20710 | 28 July 2000 (28.07.2000)   | US |
| 09/644,848     | 22 August 2000 (22.08.2000) | US |
| PCT/US00/23328 | 24 August 2000 (24.08.2000) | US |
- (60/000,000 15 September 2000 (15.09.2000) US  
PCT/US00/30952  
8 November 2000 (08.11.2000) US  
PCT/US00/32678  
1 December 2000 (01.12.2000) US  
PCT/US00/34956  
20 December 2000 (20.12.2000) US
- (71) Applicant (for all designated States except US): **GENENTECH, INC.** [US/US]; 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (72) Inventors; and  
(75) Inventors/Applicants (for US only): **BAKER, Kevin, P.** [GB/US]; 14006 Indian Run Drive, Darnestown, MD 20878 (US). **CHEN, Jian** [CN/US]; 121 York Drive, Princeton, NJ 08540 (US). **DESNOYERS, Luc** [CA/US]; 2050 Stockton Street, San Francisco, CA 94133 (US). **GODDARD, Audrey** [CA/US]; 110 Congo Street, San Francisco, CA 94131 (US). **GODOWSKI, Paul, J.** [US/US]; 2627 Easton Drive, Burlingame, CA 94010 (US). **GURNEY, Austin, L.** [US/US]; 1 Debbie Lane, Belmont, CA 94002 (US). **PAN, James** [CA/US]; 2705 Coronet Boulevard, Belmont, CA 94002 (US). **SMITH, Victoria** [AU/US]; 19 Dwight Road, Burlingame, CA 94010 (US). **WATANABE, Colin, K.** [US/US]; 128 Corliss Drive, Moraga, CA 94556 (US). **WOOD, William, I.** [US/US]; 35 Southdown Court, Hillsborough, CA 94010 (US). **ZHANG, Zemin** [CN/US]; 876 Taurus Drive, Foster City, CA 94404 (US).
- (74) Agents: **BARNES, Elizabeth, M.** et al.; c/o Genentech, Inc., MS49, 1 DNA Way, South San Francisco, CA 94080-4990 (US).
- (81) Designated States (national): AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CR, CU, CZ, DE, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) Designated States (regional): ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

[Continued on next page]

(54) Title: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

(57) Abstract: The present invention is directed to novel polypeptides and to nucleic acid molecules encoding those polypeptides. Also provided herein are vectors and host cells comprising those nucleic acid sequences, chimeric polypeptide molecules comprising the polypeptides of the present invention fused to heterologous polypeptide sequences, antibodies which bind to the polypeptides of the present invention and to methods for producing the polypeptides of the present invention.

WO 01/68848 A2



**Published:**

— without international search report and to be republished  
upon receipt of that report

*For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.*

## SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC ACIDS ENCODING THE SAME

### FIELD OF THE INVENTION

5 The present invention relates generally to the identification and isolation of novel DNA and to the recombinant production of novel polypeptides.

### BACKGROUND OF THE INVENTION

10 Extracellular proteins play important roles in, among other things, the formation, differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins. These secreted polypeptides or signaling molecules normally pass through the cellular secretory pathway to reach their site of  
15 action in the extracellular environment.

Secreted proteins have various industrial applications, including as pharmaceuticals, diagnostics, biosensors and bioreactors. Most protein drugs available at present, such as thrombolytic agents, interferons, interleukins, erythropoietins, colony stimulating factors, and various other cytokines, are secretory proteins. Their receptors, which are membrane proteins, also have potential as therapeutic or diagnostic agents. Efforts  
20 are being undertaken by both industry and academia to identify new, native secreted proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel secreted proteins. Examples of screening methods and techniques are described in the literature [see, for example, Klein et al., *Proc. Natl. Acad. Sci.* 93:7108-7113 (1996); U.S. Patent No. 5,536,637].

Membrane-bound proteins and receptors can play important roles in, among other things, the formation,  
25 differentiation and maintenance of multicellular organisms. The fate of many individual cells, e.g., proliferation, migration, differentiation, or interaction with other cells, is typically governed by information received from other cells and/or the immediate environment. This information is often transmitted by secreted polypeptides (for instance, mitogenic factors, survival factors, cytotoxic factors, differentiation factors, neuropeptides, and hormones) which are, in turn, received and interpreted by diverse cell receptors or membrane-bound proteins.  
30 Such membrane-bound proteins and cell receptors include, but are not limited to, cytokine receptors, receptor kinases, receptor phosphatases, receptors involved in cell-cell interactions, and cellular adhesion molecules like selectins and integrins. For instance, transduction of signals that regulate cell growth and differentiation is regulated in part by phosphorylation of various cellular proteins. Protein tyrosine kinases, enzymes that catalyze that process, can also act as growth factor receptors. Examples include fibroblast growth factor receptor and

nerve growth factor receptor.

Membrane-bound proteins and receptor molecules have various industrial applications, including as pharmaceutical and diagnostic agents. Receptor immunoadhesins, for instance, can be employed as therapeutic agents to block receptor-ligand interactions. The membrane-bound proteins can also be employed for screening of potential peptide or small molecule inhibitors of the relevant receptor/ligand interaction.

5        Efforts are being undertaken by both industry and academia to identify new, native receptor or membrane-bound proteins. Many efforts are focused on the screening of mammalian recombinant DNA libraries to identify the coding sequences for novel receptor or membrane-bound proteins.

#### SUMMARY OF THE INVENTION

10        In one embodiment, the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence that encodes a PRO polypeptide.

15        In one aspect, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity to (a) a DNA molecule encoding a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

25        In other aspects, the isolated nucleic acid molecule comprises a nucleotide sequence having at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94%  
30  
35



nucleic acid sequence identity, alternatively at least about 95 % nucleic acid sequence identity, alternatively at least about 96 % nucleic acid sequence identity, alternatively at least about 97 % nucleic acid sequence identity, alternatively at least about 98 % nucleic acid sequence identity and alternatively at least about 99 % nucleic acid sequence identity to (a) a DNA molecule comprising the coding sequence of a full-length PRO polypeptide cDNA as disclosed herein, the coding sequence of a PRO polypeptide lacking the signal peptide as disclosed herein, the coding sequence of an extracellular domain of a transmembrane PRO polypeptide, with or without the signal peptide, as disclosed herein or the coding sequence of any other specifically defined fragment of the full-length amino acid sequence as disclosed herein, or (b) the complement of the DNA molecule of (a).

In a further aspect, the invention concerns an isolated nucleic acid molecule comprising a nucleotide sequence having at least about 80 % nucleic acid sequence identity, alternatively at least about 81 % nucleic acid sequence identity, alternatively at least about 82 % nucleic acid sequence identity, alternatively at least about 83 % nucleic acid sequence identity, alternatively at least about 84 % nucleic acid sequence identity, alternatively at least about 85 % nucleic acid sequence identity, alternatively at least about 86 % nucleic acid sequence identity, alternatively at least about 87 % nucleic acid sequence identity, alternatively at least about 88 % nucleic acid sequence identity, alternatively at least about 89 % nucleic acid sequence identity, alternatively at least about 90 % nucleic acid sequence identity, alternatively at least about 91 % nucleic acid sequence identity, alternatively at least about 92 % nucleic acid sequence identity, alternatively at least about 93 % nucleic acid sequence identity, alternatively at least about 94 % nucleic acid sequence identity, alternatively at least about 95 % nucleic acid sequence identity, alternatively at least about 96 % nucleic acid sequence identity, alternatively at least about 97 % nucleic acid sequence identity, alternatively at least about 98 % nucleic acid sequence identity and alternatively at least about 99 % nucleic acid sequence identity to (a) a DNA molecule that encodes the same mature polypeptide encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein, or (b) the complement of the DNA molecule of (a).

Another aspect the invention provides an isolated nucleic acid molecule comprising a nucleotide sequence encoding a PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated, or is complementary to such encoding nucleotide sequence, wherein the transmembrane domain(s) of such polypeptide are disclosed herein. Therefore, soluble extracellular domains of the herein described PRO polypeptides are contemplated.

Another embodiment is directed to fragments of a PRO polypeptide coding sequence, or the complement thereof, that may find use as, for example, hybridization probes, for encoding fragments of a PRO polypeptide that may optionally encode a polypeptide comprising a binding site for an anti-PRO antibody or as antisense oligonucleotide probes. Such nucleic acid fragments are usually at least about 10 nucleotides in length, alternatively at least about 15 nucleotides in length, alternatively at least about 20 nucleotides in length, alternatively at least about 30 nucleotides in length, alternatively at least about 40 nucleotides in length, alternatively at least about 50 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 70 nucleotides in length, alternatively at least about 80 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 100 nucleotides in length, alternatively at least about 110 nucleotides in length, alternatively at least about 120 nucleotides in length,

alternatively at least about 130 nucleotides in length, alternatively at least about 140 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 160 nucleotides in length, alternatively at least about 170 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 190 nucleotides in length, alternatively at least about 200 nucleotides in length, alternatively at least about 250 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 350 nucleotides in length, alternatively at least about 400 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 500 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 700 nucleotides in length, alternatively at least about 800 nucleotides in length, alternatively at least about 900 nucleotides in length and alternatively at least about 1000 nucleotides in length, wherein in this context the term "about" means the referenced nucleotide sequence length plus or minus 10% of that referenced length. It is noted that novel fragments of a PRO polypeptide-encoding nucleotide sequence may be determined in a routine manner by aligning the PRO polypeptide-encoding nucleotide sequence with other known nucleotide sequences using any of a number of well known sequence alignment programs and determining which PRO polypeptide-encoding nucleotide sequence fragment(s) are novel. All of such PRO polypeptide-encoding nucleotide sequences are contemplated herein. Also contemplated are the PRO polypeptide fragments encoded by these nucleotide molecule fragments, preferably those PRO polypeptide fragments that comprise a binding site for an anti-PRO antibody.

In another embodiment, the invention provides isolated PRO polypeptide encoded by any of the isolated nucleic acid sequences hereinabove identified.

In a certain aspect, the invention concerns an isolated PRO polypeptide, comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a PRO polypeptide having a full-length amino acid sequence as disclosed herein, an amino acid sequence lacking the signal peptide as disclosed herein, an extracellular domain of a transmembrane protein, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of the full-length amino acid sequence as disclosed herein.

In a further aspect, the invention concerns an isolated PRO polypeptide comprising an amino acid sequence having at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83%

amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to an amino acid sequence encoded by any of the human protein cDNAs deposited with the ATCC as disclosed herein.

In a specific aspect, the invention provides an isolated PRO polypeptide without the N-terminal signal sequence and/or the initiating methionine and is encoded by a nucleotide sequence that encodes such an amino acid sequence as hereinbefore described. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

Another aspect the invention provides an isolated PRO polypeptide which is either transmembrane domain-deleted or transmembrane domain-inactivated. Processes for producing the same are also herein described, wherein those processes comprise culturing a host cell comprising a vector which comprises the appropriate encoding nucleic acid molecule under conditions suitable for expression of the PRO polypeptide and recovering the PRO polypeptide from the cell culture.

In yet another embodiment, the invention concerns agonists and antagonists of a native PRO polypeptide as defined herein. In a particular embodiment, the agonist or antagonist is an anti-PRO antibody or a small molecule.

In a further embodiment, the invention concerns a method of identifying agonists or antagonists to a PRO polypeptide which comprise contacting the PRO polypeptide with a candidate molecule and monitoring a biological activity mediated by said PRO polypeptide. Preferably, the PRO polypeptide is a native PRO polypeptide.

In a still further embodiment, the invention concerns a composition of matter comprising a PRO polypeptide, or an agonist or antagonist of a PRO polypeptide as herein described, or an anti-PRO antibody, in combination with a carrier. Optionally, the carrier is a pharmaceutically acceptable carrier.

Another embodiment of the present invention is directed to the use of a PRO polypeptide, or an agonist or antagonist thereof as hereinbefore described, or an anti-PRO antibody, for the preparation of a medicament useful in the treatment of a condition which is responsive to the PRO polypeptide, an agonist or antagonist thereof or an anti-PRO antibody.

In other embodiments of the present invention, the invention provides vectors comprising DNA encoding any of the herein described polypeptides. Host cell comprising any such vector are also provided. By way of example, the host cells may be CHO cells, *E. coli*, or yeast. A process for producing any of the herein described

polypeptides is further provided and comprises culturing host cells under conditions suitable for expression of the desired polypeptide and recovering the desired polypeptide from the cell culture.

In other embodiments, the invention provides chimeric molecules comprising any of the herein described polypeptides fused to a heterologous polypeptide or amino acid sequence. Example of such chimeric molecules comprise any of the herein described polypeptides fused to an epitope tag sequence or a Fc region of an immunoglobulin.

In another embodiment, the invention provides an antibody which binds, preferably specifically, to any of the above or below described polypeptides. Optionally, the antibody is a monoclonal antibody, humanized antibody, antibody fragment or single-chain antibody.

In yet other embodiments, the invention provides oligonucleotide probes which may be useful for isolating genomic and cDNA nucleotide sequences, measuring or detecting expression of an associated gene or as antisense probes, wherein those probes may be derived from any of the above or below described nucleotide sequences. Preferred probe lengths are described above.

In yet other embodiments, the present invention is directed to methods of using the PRO polypeptides of the present invention for a variety of uses based upon the functional biological assay data presented in the Examples below.

#### BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows a nucleotide sequence (SEQ ID NO:1) of a native sequence PRO276 cDNA, wherein SEQ ID NO:1 is a clone designated herein as "DNA16435-1208".

Figure 2 shows the amino acid sequence (SEQ ID NO:2) derived from the coding sequence of SEQ ID NO:1 shown in Figure 1.

Figure 3 shows a nucleotide sequence (SEQ ID NO:3) of a native sequence PRO284 cDNA, wherein SEQ ID NO:3 is a clone designated herein as "DNA23318-1211".

Figure 4 shows the amino acid sequence (SEQ ID NO:4) derived from the coding sequence of SEQ ID NO:3 shown in Figure 3.

Figure 5 shows a nucleotide sequence (SEQ ID NO:5) of a native sequence PRO193 cDNA, wherein SEQ ID NO:5 is a clone designated herein as "DNA23322-1393".

Figure 6 shows the amino acid sequence (SEQ ID NO:6) derived from the coding sequence of SEQ ID NO:5 shown in Figure 5.

Figure 7 shows a nucleotide sequence (SEQ ID NO:7) of a native sequence PRO190 cDNA, wherein SEQ ID NO:7 is a clone designated herein as "DNA23334-1392".

Figure 8 shows the amino acid sequence (SEQ ID NO:8) derived from the coding sequence of SEQ ID NO:7 shown in Figure 7.

Figure 9 shows a nucleotide sequence (SEQ ID NO:9) of a native sequence PRO180 cDNA, wherein SEQ ID NO:9 is a clone designated herein as "DNA26843-1389".

Figure 10 shows the amino acid sequence (SEQ ID NO:10) derived from the coding sequence of SEQ ID NO:9 shown in Figure 9.

Figure 11 shows a nucleotide sequence (SEQ ID NO:11) of a native sequence PRO194 cDNA, wherein SEQ ID NO:11 is a clone designated herein as "DNA26844-1394".

Figure 12 shows the amino acid sequence (SEQ ID NO:12) derived from the coding sequence of SEQ ID NO:11 shown in Figure 11.

5 Figure 13 shows a nucleotide sequence (SEQ ID NO:13) of a native sequence PRO218 cDNA, wherein SEQ ID NO:13 is a clone designated herein as "DNA30867-1335".

Figure 14 shows the amino acid sequence (SEQ ID NO:14) derived from the coding sequence of SEQ ID NO:13 shown in Figure 13.

Figure 15 shows a nucleotide sequence (SEQ ID NO:15) of a native sequence PRO260 cDNA, wherein SEQ ID NO:15 is a clone designated herein as "DNA33470-1175".

10 Figure 16 shows the amino acid sequence (SEQ ID NO:16) derived from the coding sequence of SEQ ID NO:15 shown in Figure 15.

Figure 17 shows a nucleotide sequence (SEQ ID NO:17) of a native sequence PRO233 cDNA, wherein SEQ ID NO:17 is a clone designated herein as "DNA34436-1238".

15 Figure 18 shows the amino acid sequence (SEQ ID NO:18) derived from the coding sequence of SEQ ID NO:17 shown in Figure 17.

Figure 19 shows a nucleotide sequence (SEQ ID NO:19) of a native sequence PRO234 cDNA, wherein SEQ ID NO:19 is a clone designated herein as "DNA35557-1137".

Figure 20 shows the amino acid sequence (SEQ ID NO:20) derived from the coding sequence of SEQ ID NO:19 shown in Figure 19.

20 Figure 21 shows a nucleotide sequence (SEQ ID NO:21) of a native sequence PRO236 cDNA, wherein SEQ ID NO:21 is a clone designated herein as "DNA35599-1168".

Figure 22 shows the amino acid sequence (SEQ ID NO:22) derived from the coding sequence of SEQ ID NO:21 shown in Figure 21.

25 Figure 23 shows a nucleotide sequence (SEQ ID NO:23) of a native sequence PRO244 cDNA, wherein SEQ ID NO:23 is a clone designated herein as "DNA35668-1171".

Figure 24 shows the amino acid sequence (SEQ ID NO:24) derived from the coding sequence of SEQ ID NO:23 shown in Figure 23.

Figure 25 shows a nucleotide sequence (SEQ ID NO:25) of a native sequence PRO262 cDNA, wherein SEQ ID NO:25 is a clone designated herein as "DNA36992-1168".

30 Figure 26 shows the amino acid sequence (SEQ ID NO:26) derived from the coding sequence of SEQ ID NO:25 shown in Figure 25.

Figure 27 shows a nucleotide sequence (SEQ ID NO:27) of a native sequence PRO271 cDNA, wherein SEQ ID NO:27 is a clone designated herein as "DNA39423-1182".

35 Figure 28 shows the amino acid sequence (SEQ ID NO:28) derived from the coding sequence of SEQ ID NO:27 shown in Figure 27.

Figure 29 shows a nucleotide sequence (SEQ ID NO:29) of a native sequence PRO268 cDNA, wherein SEQ ID NO:29 is a clone designated herein as "DNA39427-1179".

Figure 30 shows the amino acid sequence (SEQ ID NO:30) derived from the coding sequence of SEQ ID NO:29 shown in Figure 29.

Figure 31 shows a nucleotide sequence (SEQ ID NO:31) of a native sequence PRO270 cDNA, wherein SEQ ID NO:31 is a clone designated herein as "DNA39510-1181".

5     Figure 32 shows the amino acid sequence (SEQ ID NO:32) derived from the coding sequence of SEQ ID NO:31 shown in Figure 31.

Figure 33 shows a nucleotide sequence (SEQ ID NO:33) of a native sequence PRO355 cDNA, wherein SEQ ID NO:33 is a clone designated herein as "DNA39518-1247".

Figure 34 shows the amino acid sequence (SEQ ID NO:34) derived from the coding sequence of SEQ ID NO:33 shown in Figure 33.

10     Figure 35 shows a nucleotide sequence (SEQ ID NO:35) of a native sequence PRO298 cDNA, wherein SEQ ID NO:35 is a clone designated herein as "DNA39975-1210".

Figure 36 shows the amino acid sequence (SEQ ID NO:36) derived from the coding sequence of SEQ ID NO:35 shown in Figure 35.

15     Figure 37 shows a nucleotide sequence (SEQ ID NO:37) of a native sequence PRO299 cDNA, wherein SEQ ID NO:37 is a clone designated herein as "DNA39976-1215".

Figure 38 shows the amino acid sequence (SEQ ID NO:38) derived from the coding sequence of SEQ ID NO:37 shown in Figure 37.

Figure 39 shows a nucleotide sequence (SEQ ID NO:39) of a native sequence PRO296 cDNA, wherein SEQ ID NO:39 is a clone designated herein as "DNA39979-1213".

20     Figure 40 shows the amino acid sequence (SEQ ID NO:40) derived from the coding sequence of SEQ ID NO:39 shown in Figure 39.

Figure 41 shows a nucleotide sequence (SEQ ID NO:41) of a native sequence PRO329 cDNA, wherein SEQ ID NO:41 is a clone designated herein as "DNA40594-1233".

25     Figure 42 shows the amino acid sequence (SEQ ID NO:42) derived from the coding sequence of SEQ ID NO:41 shown in Figure 41.

Figure 43 shows a nucleotide sequence (SEQ ID NO:43) of a native sequence PRO330 cDNA, wherein SEQ ID NO:43 is a clone designated herein as "DNA40603-1232".

Figure 44 shows the amino acid sequence (SEQ ID NO:44) derived from the coding sequence of SEQ ID NO:43 shown in Figure 43.

30     Figure 45 shows a nucleotide sequence (SEQ ID NO:45) of a native sequence PRO294 cDNA, wherein SEQ ID NO:45 is a clone designated herein as "DNA40604-1187".

Figure 46 shows the amino acid sequence (SEQ ID NO:46) derived from the coding sequence of SEQ ID NO:45 shown in Figure 45.

35     Figure 47 shows a nucleotide sequence (SEQ ID NO:47) of a native sequence PRO300 cDNA, wherein SEQ ID NO:47 is a clone designated herein as "DNA40625-1189".

Figure 48 shows the amino acid sequence (SEQ ID NO:48) derived from the coding sequence of SEQ ID NO:47 shown in Figure 47.

Figure 49 shows a nucleotide sequence (SEQ ID NO:49) of a native sequence PRO307 cDNA, wherein SEQ ID NO:49 is a clone designated herein as "DNA41225-1217".

Figure 50 shows the amino acid sequence (SEQ ID NO:50) derived from the coding sequence of SEQ ID NO:49 shown in Figure 49.

5 Figure 51 shows a nucleotide sequence (SEQ ID NO:51) of a native sequence PRO334 cDNA, wherein SEQ ID NO:51 is a clone designated herein as "DNA41379-1236".

Figure 52 shows the amino acid sequence (SEQ ID NO:52) derived from the coding sequence of SEQ ID NO:51 shown in Figure 51.

Figure 53 shows a nucleotide sequence (SEQ ID NO:53) of a native sequence PRO352 cDNA, wherein SEQ ID NO:53 is a clone designated herein as "DNA41386-1316".

10 Figure 54 shows the amino acid sequence (SEQ ID NO:54) derived from the coding sequence of SEQ ID NO:53 shown in Figure 53.

Figure 55 shows a nucleotide sequence (SEQ ID NO:55) of a native sequence PRO710 cDNA, wherein SEQ ID NO:55 is a clone designated herein as "DNA44161-1434".

15 Figure 56 shows the amino acid sequence (SEQ ID NO:56) derived from the coding sequence of SEQ ID NO:55 shown in Figure 55.

Figure 57 shows a nucleotide sequence (SEQ ID NO:57) of a native sequence PRO873 cDNA, wherein SEQ ID NO:57 is a clone designated herein as "DNA44179-1362".

Figure 58 shows the amino acid sequence (SEQ ID NO:58) derived from the coding sequence of SEQ ID NO:57 shown in Figure 57.

20 Figure 59 shows a nucleotide sequence (SEQ ID NO:59) of a native sequence PRO354 cDNA, wherein SEQ ID NO:59 is a clone designated herein as "DNA44192-1246".

Figure 60 shows the amino acid sequence (SEQ ID NO:60) derived from the coding sequence of SEQ ID NO:59 shown in Figure 59.

25 Figure 61 shows a nucleotide sequence (SEQ ID NO:61) of a native sequence PRO1151 cDNA, wherein SEQ ID NO:61 is a clone designated herein as "DNA44694-1500".

Figure 62 shows the amino acid sequence (SEQ ID NO:62) derived from the coding sequence of SEQ ID NO:61 shown in Figure 61.

Figure 63 shows a nucleotide sequence (SEQ ID NO:63) of a native sequence PRO382 cDNA, wherein SEQ ID NO:63 is a clone designated herein as "DNA45234-1277".

30 Figure 64 shows the amino acid sequence (SEQ ID NO:64) derived from the coding sequence of SEQ ID NO:63 shown in Figure 63.

Figure 65 shows a nucleotide sequence (SEQ ID NO:65) of a native sequence PRO1864 cDNA, wherein SEQ ID NO:65 is a clone designated herein as "DNA45409-2511".

35 Figure 66 shows the amino acid sequence (SEQ ID NO:66) derived from the coding sequence of SEQ ID NO:65 shown in Figure 65.

Figure 67 shows a nucleotide sequence (SEQ ID NO:67) of a native sequence PRO386 cDNA, wherein SEQ ID NO:67 is a clone designated herein as "DNA45415-1318".

Figure 68 shows the amino acid sequence (SEQ ID NO:68) derived from the coding sequence of SEQ ID NO:67 shown in Figure 67.

Figure 69 shows a nucleotide sequence (SEQ ID NO:69) of a native sequence PRO541 cDNA, wherein SEQ ID NO:69 is a clone designated herein as "DNA45417-1432".

5 Figure 70 shows the amino acid sequence (SEQ ID NO:70) derived from the coding sequence of SEQ ID NO:69 shown in Figure 69.

Figure 71 shows a nucleotide sequence (SEQ ID NO:71) of a native sequence PRO852 cDNA, wherein SEQ ID NO:71 is a clone designated herein as "DNA45493-1349".

Figure 72 shows the amino acid sequence (SEQ ID NO:72) derived from the coding sequence of SEQ ID NO:71 shown in Figure 71.

10 Figure 73 shows a nucleotide sequence (SEQ ID NO:73) of a native sequence PRO700 cDNA, wherein SEQ ID NO:73 is a clone designated herein as "DNA46776-1284".

Figure 74 shows the amino acid sequence (SEQ ID NO:74) derived from the coding sequence of SEQ ID NO:73 shown in Figure 73.

15 Figures 75A-75B show a nucleotide sequence (SEQ ID NO:75) of a native sequence PRO708 cDNA, wherein SEQ ID NO:75 is a clone designated herein as "DNA48296-1292".

Figure 76 shows the amino acid sequence (SEQ ID NO:76) derived from the coding sequence of SEQ ID NO:75 shown in Figures 75A-75B.

Figure 77 shows a nucleotide sequence (SEQ ID NO:77) of a native sequence PRO707 cDNA, wherein SEQ ID NO:77 is a clone designated herein as "DNA48306-1291".

20 Figure 78 shows the amino acid sequence (SEQ ID NO:78) derived from the coding sequence of SEQ ID NO:77 shown in Figure 77.

Figure 79 shows a nucleotide sequence (SEQ ID NO:79) of a native sequence PRO864 cDNA, wherein SEQ ID NO:79 is a clone designated herein as "DNA48328-1355".

25 Figure 80 shows the amino acid sequence (SEQ ID NO:80) derived from the coding sequence of SEQ ID NO:79 shown in Figure 79.

Figure 81 shows a nucleotide sequence (SEQ ID NO:81) of a native sequence PRO706 cDNA, wherein SEQ ID NO:81 is a clone designated herein as "DNA48329-1290".

Figure 82 shows the amino acid sequence (SEQ ID NO:82) derived from the coding sequence of SEQ ID NO:81 shown in Figure 81.

30 Figure 83 shows a nucleotide sequence (SEQ ID NO:83) of a native sequence PRO732 cDNA, wherein SEQ ID NO:83 is a clone designated herein as "DNA48334-1435".

Figure 84 shows the amino acid sequence (SEQ ID NO:84) derived from the coding sequence of SEQ ID NO:83 shown in Figure 83.

35 Figure 85 shows a nucleotide sequence (SEQ ID NO:85) of a native sequence PRO537 cDNA, wherein SEQ ID NO:85 is a clone designated herein as "DNA49141-1431".

Figure 86 shows the amino acid sequence (SEQ ID NO:86) derived from the coding sequence of SEQ ID NO:85 shown in Figure 85.



Figure 87 shows a nucleotide sequence (SEQ ID NO:87) of a native sequence PRO545 cDNA, wherein SEQ ID NO:87 is a clone designated herein as "DNA49624-1279".

Figure 88 shows the amino acid sequence (SEQ ID NO:88) derived from the coding sequence of SEQ ID NO:87 shown in Figure 87.

5 Figure 89 shows a nucleotide sequence (SEQ ID NO:89) of a native sequence PRO718 cDNA, wherein SEQ ID NO:89 is a clone designated herein as "DNA49647-1398".

Figure 90 shows the amino acid sequence (SEQ ID NO:90) derived from the coding sequence of SEQ ID NO:89 shown in Figure 89.

Figure 91 shows a nucleotide sequence (SEQ ID NO:91) of a native sequence PRO872 cDNA, wherein SEQ ID NO:91 is a clone designated herein as "DNA49819-1439".

10 Figure 92 shows the amino acid sequence (SEQ ID NO:92) derived from the coding sequence of SEQ ID NO:91 shown in Figure 91.

Figure 93 shows a nucleotide sequence (SEQ ID NO:93) of a native sequence PRO704 cDNA, wherein SEQ ID NO:93 is a clone designated herein as "DNA50911-1288".

15 Figure 94 shows the amino acid sequence (SEQ ID NO:94) derived from the coding sequence of SEQ ID NO:93 shown in Figure 93.

Figure 95 shows a nucleotide sequence (SEQ ID NO:95) of a native sequence PRO705 cDNA, wherein SEQ ID NO:95 is a clone designated herein as "DNA50914-1289".

Figure 96 shows the amino acid sequence (SEQ ID NO:96) derived from the coding sequence of SEQ ID NO:95 shown in Figure 95.

20 Figure 97 shows a nucleotide sequence (SEQ ID NO:97) of a native sequence PRO871 cDNA, wherein SEQ ID NO:97 is a clone designated herein as "DNA50919-1361".

Figure 98 shows the amino acid sequence (SEQ ID NO:98) derived from the coding sequence of SEQ ID NO:97 shown in Figure 97.

25 Figure 99 shows a nucleotide sequence (SEQ ID NO:99) of a native sequence PRO702 cDNA, wherein SEQ ID NO:99 is a clone designated herein as "DNA50980-1286".

Figure 100 shows the amino acid sequence (SEQ ID NO:100) derived from the coding sequence of SEQ ID NO:99 shown in Figure 99.

Figure 101 shows a nucleotide sequence (SEQ ID NO:101) of a native sequence PRO944 cDNA, wherein SEQ ID NO:101 is a clone designated herein as "DNA52185-1370".

30 Figure 102 shows the amino acid sequence (SEQ ID NO:102) derived from the coding sequence of SEQ ID NO:101 shown in Figure 101.

Figure 103 shows a nucleotide sequence (SEQ ID NO:103) of a native sequence PRO739 cDNA, wherein SEQ ID NO:103 is a clone designated herein as "DNA52756".

35 Figure 104 shows the amino acid sequence (SEQ ID NO:104) derived from the coding sequence of SEQ ID NO:103 shown in Figure 103.

Figure 105 shows a nucleotide sequence (SEQ ID NO:105) of a native sequence PRO941 cDNA, wherein SEQ ID NO:105 is a clone designated herein as "DNA53906-1368".

Figure 106 shows the amino acid sequence (SEQ ID NO:106) derived from the coding sequence of SEQ ID NO:105 shown in Figure 105.

Figure 107 shows a nucleotide sequence (SEQ ID NO:107) of a native sequence PRO1082 cDNA, wherein SEQ ID NO:107 is a clone designated herein as "DNA53912-1457".

5 Figure 108 shows the amino acid sequence (SEQ ID NO:108) derived from the coding sequence of SEQ ID NO:107 shown in Figure 107.

Figure 109 shows a nucleotide sequence (SEQ ID NO:109) of a native sequence PRO1133 cDNA, wherein SEQ ID NO:109 is a clone designated herein as "DNA53913-1490".

Figure 110 shows the amino acid sequence (SEQ ID NO:110) derived from the coding sequence of SEQ ID NO:109 shown in Figure 109.

10 Figure 111 shows a nucleotide sequence (SEQ ID NO:111) of a native sequence PRO983 cDNA, wherein SEQ ID NO:111 is a clone designated herein as "DNA53977-1371".

Figure 112 shows the amino acid sequence (SEQ ID NO:112) derived from the coding sequence of SEQ ID NO:111 shown in Figure 111.

15 Figure 113 shows a nucleotide sequence (SEQ ID NO:113) of a native sequence PRO784 cDNA, wherein SEQ ID NO:113 is a clone designated herein as "DNA53978-1443".

Figure 114 shows the amino acid sequence (SEQ ID NO:114) derived from the coding sequence of SEQ ID NO:113 shown in Figure 113.

Figure 115 shows a nucleotide sequence (SEQ ID NO:115) of a native sequence PRO783 cDNA, wherein SEQ ID NO:115 is a clone designated herein as "DNA53996-1442".

20 Figure 116 shows the amino acid sequence (SEQ ID NO:116) derived from the coding sequence of SEQ ID NO:115 shown in Figure 115.

Figure 117 shows a nucleotide sequence (SEQ ID NO:117) of a native sequence PRO940 cDNA, wherein SEQ ID NO:117 is a clone designated herein as "DNA54002-1367".

25 Figure 118 shows the amino acid sequence (SEQ ID NO:118) derived from the coding sequence of SEQ ID NO:117 shown in Figure 117.

Figure 119 shows a nucleotide sequence (SEQ ID NO:119) of a native sequence PRO768 cDNA, wherein SEQ ID NO:119 is a clone designated herein as "DNA55737-1345".

Figure 120 shows the amino acid sequence (SEQ ID NO:120) derived from the coding sequence of SEQ ID NO:119 shown in Figure 119.

30 Figure 121 shows a nucleotide sequence (SEQ ID NO:121) of a native sequence PRO1079 cDNA, wherein SEQ ID NO:121 is a clone designated herein as "DNA56050-1455".

Figure 122 shows the amino acid sequence (SEQ ID NO:122) derived from the coding sequence of SEQ ID NO:121 shown in Figure 121.

35 Figure 123 shows a nucleotide sequence (SEQ ID NO:123) of a native sequence PRO1078 cDNA, wherein SEQ ID NO:123 is a clone designated herein as "DNA56052-1454".

Figure 124 shows the amino acid sequence (SEQ ID NO:124) derived from the coding sequence of SEQ ID NO:123 shown in Figure 123.

Figure 125 shows a nucleotide sequence (SEQ ID NO:125) of a native sequence PRO1018 cDNA, wherein SEQ ID NO:125 is a clone designated herein as "DNA56107-1415".

Figure 126 shows the amino acid sequence (SEQ ID NO:126) derived from the coding sequence of SEQ ID NO:125 shown in Figure 125.

5 Figure 127 shows a nucleotide sequence (SEQ ID NO:127) of a native sequence PRO793 cDNA, wherein SEQ ID NO:127 is a clone designated herein as "DNA56110-1437".

Figure 128 shows the amino acid sequence (SEQ ID NO:128) derived from the coding sequence of SEQ ID NO:127 shown in Figure 127.

Figure 129 shows a nucleotide sequence (SEQ ID NO:129) of a native sequence PRO1773 cDNA, wherein SEQ ID NO:129 is a clone designated herein as "DNA56406-1704".

10 Figure 130 shows the amino acid sequence (SEQ ID NO:130) derived from the coding sequence of SEQ ID NO:129 shown in Figure 129.

Figure 131 shows a nucleotide sequence (SEQ ID NO:131) of a native sequence PRO1014 cDNA, wherein SEQ ID NO:131 is a clone designated herein as "DNA56409-1377".

15 Figure 132 shows the amino acid sequence (SEQ ID NO:132) derived from the coding sequence of SEQ ID NO:131 shown in Figure 131.

Figure 133 shows a nucleotide sequence (SEQ ID NO:133) of a native sequence PRO1013 cDNA, wherein SEQ ID NO:133 is a clone designated herein as "DNA56410-1414".

Figure 134 shows the amino acid sequence (SEQ ID NO:134) derived from the coding sequence of SEQ ID NO:133 shown in Figure 133.

20 Figure 135 shows a nucleotide sequence (SEQ ID NO:135) of a native sequence PRO937 cDNA, wherein SEQ ID NO:135 is a clone designated herein as "DNA56436-1448".

Figure 136 shows the amino acid sequence (SEQ ID NO:136) derived from the coding sequence of SEQ ID NO:135 shown in Figure 135.

25 Figure 137 shows a nucleotide sequence (SEQ ID NO:137) of a native sequence PRO1477 cDNA, wherein SEQ ID NO:137 is a clone designated herein as "DNA56529-1647".

Figure 138 shows the amino acid sequence (SEQ ID NO:138) derived from the coding sequence of SEQ ID NO:137 shown in Figure 137.

Figure 139 shows a nucleotide sequence (SEQ ID NO:139) of a native sequence PRO842 cDNA, wherein SEQ ID NO:139 is a clone designated herein as "DNA56855-1447".

30 Figure 140 shows the amino acid sequence (SEQ ID NO:140) derived from the coding sequence of SEQ ID NO:139 shown in Figure 139.

Figure 141 shows a nucleotide sequence (SEQ ID NO:141) of a native sequence PRO839 cDNA, wherein SEQ ID NO:141 is a clone designated herein as "DNA56859-1445".

35 Figure 142 shows the amino acid sequence (SEQ ID NO:142) derived from the coding sequence of SEQ ID NO:141 shown in Figure 141.

Figure 143 shows a nucleotide sequence (SEQ ID NO:143) of a native sequence PRO1180 cDNA, wherein SEQ ID NO:143 is a clone designated herein as "DNA56860-1510".

Figure 144 shows the amino acid sequence (SEQ ID NO:144) derived from the coding sequence of SEQ ID NO:143 shown in Figure 143.

Figure 145 shows a nucleotide sequence (SEQ ID NO:145) of a native sequence PRO1134 cDNA, wherein SEQ ID NO:145 is a clone designated herein as "DNA56865-1491".

5 Figure 146 shows the amino acid sequence (SEQ ID NO:146) derived from the coding sequence of SEQ ID NO:145 shown in Figure 145.

Figure 147 shows a nucleotide sequence (SEQ ID NO:147) of a native sequence PRO1115 cDNA, wherein SEQ ID NO:147 is a clone designated herein as "DNA56868-1478".

Figure 148 shows the amino acid sequence (SEQ ID NO:148) derived from the coding sequence of SEQ ID NO:147 shown in Figure 147.

10 Figure 149 shows a nucleotide sequence (SEQ ID NO:149) of a native sequence PRO1277 cDNA, wherein SEQ ID NO:149 is a clone designated herein as "DNA56869-1545".

Figure 150 shows the amino acid sequence (SEQ ID NO:150) derived from the coding sequence of SEQ ID NO:149 shown in Figure 149.

15 Figure 151 shows a nucleotide sequence (SEQ ID NO:151) of a native sequence PRO1135 cDNA, wherein SEQ ID NO:151 is a clone designated herein as "DNA56870-1492".

Figure 152 shows the amino acid sequence (SEQ ID NO:152) derived from the coding sequence of SEQ ID NO:151 shown in Figure 151.

Figure 153 shows a nucleotide sequence (SEQ ID NO:153) of a native sequence PRO827 cDNA, wherein SEQ ID NO:153 is a clone designated herein as "DNA57039-1402".

20 Figure 154 shows the amino acid sequence (SEQ ID NO:154) derived from the coding sequence of SEQ ID NO:153 shown in Figure 153.

Figure 155 shows a nucleotide sequence (SEQ ID NO:155) of a native sequence PRO1057 cDNA, wherein SEQ ID NO:155 is a clone designated herein as "DNA57253-1382".

25 Figure 156 shows the amino acid sequence (SEQ ID NO:156) derived from the coding sequence of SEQ ID NO:155 shown in Figure 155.

Figure 157 shows a nucleotide sequence (SEQ ID NO:157) of a native sequence PRO1113 cDNA, wherein SEQ ID NO:157 is a clone designated herein as "DNA57254-1477".

Figure 158 shows the amino acid sequence (SEQ ID NO:158) derived from the coding sequence of SEQ ID NO:157 shown in Figure 157.

30 Figure 159 shows a nucleotide sequence (SEQ ID NO:159) of a native sequence PRO1006 cDNA, wherein SEQ ID NO:159 is a clone designated herein as "DNA57699-1412".

Figure 160 shows the amino acid sequence (SEQ ID NO:160) derived from the coding sequence of SEQ ID NO:159 shown in Figure 159.

35 Figure 161 shows a nucleotide sequence (SEQ ID NO:161) of a native sequence PRO1074 cDNA, wherein SEQ ID NO:161 is a clone designated herein as "DNA57704-1452".

Figure 162 shows the amino acid sequence (SEQ ID NO:162) derived from the coding sequence of SEQ ID NO:161 shown in Figure 161.

Figure 163 shows a nucleotide sequence (SEQ ID NO:163) of a native sequence PRO1073 cDNA, wherein SEQ ID NO:163 is a clone designated herein as "DNA57710-1451".

Figure 164 shows the amino acid sequence (SEQ ID NO:164) derived from the coding sequence of SEQ ID NO:163 shown in Figure 163.

5 Figure 165 shows a nucleotide sequence (SEQ ID NO:165) of a native sequence PRO1136 cDNA, wherein SEQ ID NO:165 is a clone designated herein as "DNA57827-1493".

Figure 166 shows the amino acid sequence (SEQ ID NO:166) derived from the coding sequence of SEQ ID NO:165 shown in Figure 165.

Figure 167 shows a nucleotide sequence (SEQ ID NO:167) of a native sequence PRO1004 cDNA, wherein SEQ ID NO:167 is a clone designated herein as "DNA57844-1410".

10 Figure 168 shows the amino acid sequence (SEQ ID NO:168) derived from the coding sequence of SEQ ID NO:167 shown in Figure 167.

Figure 169 shows a nucleotide sequence (SEQ ID NO:169) of a native sequence PRO1344 cDNA, wherein SEQ ID NO:169 is a clone designated herein as "DNA58723-1588".

15 Figure 170 shows the amino acid sequence (SEQ ID NO:170) derived from the coding sequence of SEQ ID NO:169 shown in Figure 169.

Figure 171 shows a nucleotide sequence (SEQ ID NO:171) of a native sequence PRO1110 cDNA, wherein SEQ ID NO:171 is a clone designated herein as "DNA58727-1474".

Figure 172 shows the amino acid sequence (SEQ ID NO:172) derived from the coding sequence of SEQ ID NO:171 shown in Figure 171.

20 Figure 173 shows a nucleotide sequence (SEQ ID NO:173) of a native sequence PRO1378 cDNA, wherein SEQ ID NO:173 is a clone designated herein as "DNA58730-1607".

Figure 174 shows the amino acid sequence (SEQ ID NO:174) derived from the coding sequence of SEQ ID NO:173 shown in Figure 173.

25 Figure 175 shows a nucleotide sequence (SEQ ID NO:175) of a native sequence PRO1481 cDNA, wherein SEQ ID NO:175 is a clone designated herein as "DNA58732-1650".

Figure 176 shows the amino acid sequence (SEQ ID NO:176) derived from the coding sequence of SEQ ID NO:175 shown in Figure 175.

Figure 177 shows a nucleotide sequence (SEQ ID NO:177) of a native sequence PRO1109 cDNA, wherein SEQ ID NO:177 is a clone designated herein as "DNA58737-1473".

30 Figure 178 shows the amino acid sequence (SEQ ID NO:178) derived from the coding sequence of SEQ ID NO:177 shown in Figure 177.

Figure 179 shows a nucleotide sequence (SEQ ID NO:179) of a native sequence PRO1383 cDNA, wherein SEQ ID NO:179 is a clone designated herein as "DNA58743-1609".

35 Figure 180 shows the amino acid sequence (SEQ ID NO:180) derived from the coding sequence of SEQ ID NO:179 shown in Figure 179.

Figure 181 shows a nucleotide sequence (SEQ ID NO:181) of a native sequence PRO1072 cDNA, wherein SEQ ID NO:181 is a clone designated herein as "DNA58747-1384".

Figure 182 shows the amino acid sequence (SEQ ID NO:182) derived from the coding sequence of SEQ ID NO:181 shown in Figure 181.

Figure 183 shows a nucleotide sequence (SEQ ID NO:183) of a native sequence PRO1189 cDNA, wherein SEQ ID NO:183 is a clone designated herein as "DNA58828-1519".

5 Figure 184 shows the amino acid sequence (SEQ ID NO:184) derived from the coding sequence of SEQ ID NO:183 shown in Figure 183.

Figure 185 shows a nucleotide sequence (SEQ ID NO:185) of a native sequence PRO1003 cDNA, wherein SEQ ID NO:185 is a clone designated herein as "DNA58846-1409".

Figure 186 shows the amino acid sequence (SEQ ID NO:186) derived from the coding sequence of SEQ ID NO:185 shown in Figure 185.

10 Figure 187 shows a nucleotide sequence (SEQ ID NO:187) of a native sequence PRO1108 cDNA, wherein SEQ ID NO:187 is a clone designated herein as "DNA58848-1472".

Figure 188 shows the amino acid sequence (SEQ ID NO:188) derived from the coding sequence of SEQ ID NO:187 shown in Figure 187.

15 Figure 189 shows a nucleotide sequence (SEQ ID NO:189) of a native sequence PRO1137 cDNA, wherein SEQ ID NO:189 is a clone designated herein as "DNA58849-1494".

Figure 190 shows the amino acid sequence (SEQ ID NO:190) derived from the coding sequence of SEQ ID NO:189 shown in Figure 189.

Figure 191 shows a nucleotide sequence (SEQ ID NO:191) of a native sequence PRO1138 cDNA, wherein SEQ ID NO:191 is a clone designated herein as "DNA58850-1495".

20 Figure 192 shows the amino acid sequence (SEQ ID NO:192) derived from the coding sequence of SEQ ID NO:191 shown in Figure 191.

Figure 193 shows a nucleotide sequence (SEQ ID NO:193) of a native sequence PRO1415 cDNA, wherein SEQ ID NO:193 is a clone designated herein as "DNA58852-1637".

25 Figure 194 shows the amino acid sequence (SEQ ID NO:194) derived from the coding sequence of SEQ ID NO:193 shown in Figure 193.

Figure 195 shows a nucleotide sequence (SEQ ID NO:195) of a native sequence PRO1054 cDNA, wherein SEQ ID NO:195 is a clone designated herein as "DNA58853-1423".

Figure 196 shows the amino acid sequence (SEQ ID NO:196) derived from the coding sequence of SEQ ID NO:195 shown in Figure 195.

30 Figure 197 shows a nucleotide sequence (SEQ ID NO:197) of a native sequence PRO994 cDNA, wherein SEQ ID NO:197 is a clone designated herein as "DNA58855-1422".

Figure 198 shows the amino acid sequence (SEQ ID NO:198) derived from the coding sequence of SEQ ID NO:197 shown in Figure 197.

35 Figure 199 shows a nucleotide sequence (SEQ ID NO:199) of a native sequence PRO1069 cDNA, wherein SEQ ID NO:199 is a clone designated herein as "DNA59211-1450".

Figure 200 shows the amino acid sequence (SEQ ID NO:200) derived from the coding sequence of SEQ ID NO:199 shown in Figure 199.

Figure 201 shows a nucleotide sequence (SEQ ID NO:201) of a native sequence PRO1411 cDNA, wherein SEQ ID NO:201 is a clone designated herein as "DNA59212-1627".

Figure 202 shows the amino acid sequence (SEQ ID NO:202) derived from the coding sequence of SEQ ID NO:201 shown in Figure 201.

5 Figure 203 shows a nucleotide sequence (SEQ ID NO:203) of a native sequence PRO1129 cDNA, wherein SEQ ID NO:203 is a clone designated herein as "DNA59213-1487".

Figure 204 shows the amino acid sequence (SEQ ID NO:204) derived from the coding sequence of SEQ ID NO:203 shown in Figure 203.

Figure 205 shows a nucleotide sequence (SEQ ID NO:205) of a native sequence PRO1359 cDNA, wherein SEQ ID NO:205 is a clone designated herein as "DNA59219-1613".

10 Figure 206 shows the amino acid sequence (SEQ ID NO:206) derived from the coding sequence of SEQ ID NO:205 shown in Figure 205.

Figure 207 shows a nucleotide sequence (SEQ ID NO:207) of a native sequence PRO1139 cDNA, wherein SEQ ID NO:207 is a clone designated herein as "DNA59497-1496".

15 Figure 208 shows the amino acid sequence (SEQ ID NO:208) derived from the coding sequence of SEQ ID NO:207 shown in Figure 207.

Figure 209 shows a nucleotide sequence (SEQ ID NO:209) of a native sequence PRO1065 cDNA, wherein SEQ ID NO:209 is a clone designated herein as "DNA59602-1436".

Figure 210 shows the amino acid sequence (SEQ ID NO:210) derived from the coding sequence of SEQ ID NO:209 shown in Figure 209.

20 Figure 211 shows a nucleotide sequence (SEQ ID NO:211) of a native sequence PRO1028 cDNA, wherein SEQ ID NO:211 is a clone designated herein as "DNA59603-1419".

Figure 212 shows the amino acid sequence (SEQ ID NO:212) derived from the coding sequence of SEQ ID NO:211 shown in Figure 211.

25 Figure 213 shows a nucleotide sequence (SEQ ID NO:213) of a native sequence PRO1027 cDNA, wherein SEQ ID NO:213 is a clone designated herein as "DNA59605-1418".

Figure 214 shows the amino acid sequence (SEQ ID NO:214) derived from the coding sequence of SEQ ID NO:213 shown in Figure 213.

Figure 215 shows a nucleotide sequence (SEQ ID NO:215) of a native sequence PRO1140 cDNA, wherein SEQ ID NO:215 is a clone designated herein as "DNA59607-1497".

30 Figure 216 shows the amino acid sequence (SEQ ID NO:216) derived from the coding sequence of SEQ ID NO:215 shown in Figure 215.

Figure 217 shows a nucleotide sequence (SEQ ID NO:217) of a native sequence PRO1291 cDNA, wherein SEQ ID NO:217 is a clone designated herein as "DNA59610-1556".

35 Figure 218 shows the amino acid sequence (SEQ ID NO:218) derived from the coding sequence of SEQ ID NO:217 shown in Figure 217.

Figure 219 shows a nucleotide sequence (SEQ ID NO:219) of a native sequence PRO1105 cDNA, wherein SEQ ID NO:219 is a clone designated herein as "DNA59612-1466".

Figure 220 shows the amino acid sequence (SEQ ID NO:220) derived from the coding sequence of SEQ ID NO:219 shown in Figure 219.

Figure 221 shows a nucleotide sequence (SEQ ID NO:221) of a native sequence PRO1026 cDNA, wherein SEQ ID NO:221 is a clone designated herein as "DNA59613-1417".

5 Figure 222 shows the amino acid sequence (SEQ ID NO:222) derived from the coding sequence of SEQ ID NO:221 shown in Figure 221.

Figure 223 shows a nucleotide sequence (SEQ ID NO:223) of a native sequence PRO1104 cDNA, wherein SEQ ID NO:223 is a clone designated herein as "DNA59616-1465".

Figure 224 shows the amino acid sequence (SEQ ID NO:224) derived from the coding sequence of SEQ ID NO:223 shown in Figure 223.

10 Figure 225 shows a nucleotide sequence (SEQ ID NO:225) of a native sequence PRO1100 cDNA, wherein SEQ ID NO:225 is a clone designated herein as "DNA59619-1464".

Figure 226 shows the amino acid sequence (SEQ ID NO:226) derived from the coding sequence of SEQ ID NO:225 shown in Figure 225.

15 Figure 227 shows a nucleotide sequence (SEQ ID NO:227) of a native sequence PRO1141 cDNA, wherein SEQ ID NO:227 is a clone designated herein as "DNA59625-1498".

Figure 228 shows the amino acid sequence (SEQ ID NO:228) derived from the coding sequence of SEQ ID NO:227 shown in Figure 227.

Figure 229 shows a nucleotide sequence (SEQ ID NO:229) of a native sequence PRO1772 cDNA, wherein SEQ ID NO:229 is a clone designated herein as "DNA59817-1703".

20 Figure 230 shows the amino acid sequence (SEQ ID NO:230) derived from the coding sequence of SEQ ID NO:229 shown in Figure 229.

Figure 231 shows a nucleotide sequence (SEQ ID NO:231) of a native sequence PRO1064 cDNA, wherein SEQ ID NO:231 is a clone designated herein as "DNA59827-1426".

25 Figure 232 shows the amino acid sequence (SEQ ID NO:232) derived from the coding sequence of SEQ ID NO:231 shown in Figure 231.

Figure 233 shows a nucleotide sequence (SEQ ID NO:233) of a native sequence PRO1379 cDNA, wherein SEQ ID NO:233 is a clone designated herein as "DNA59828-1608".

Figure 234 shows the amino acid sequence (SEQ ID NO:234) derived from the coding sequence of SEQ ID NO:233 shown in Figure 233.

30 Figure 235 shows a nucleotide sequence (SEQ ID NO:235) of a native sequence PRO3573 cDNA, wherein SEQ ID NO:235 is a clone designated herein as "DNA59837-2545".

Figure 236 shows the amino acid sequence (SEQ ID NO:236) derived from the coding sequence of SEQ ID NO:235 shown in Figure 235.

35 Figure 237 shows a nucleotide sequence (SEQ ID NO:237) of a native sequence PRO3566 cDNA, wherein SEQ ID NO:237 is a clone designated herein as "DNA59844-2542".

Figure 238 shows the amino acid sequence (SEQ ID NO:238) derived from the coding sequence of SEQ ID NO:237 shown in Figure 237.



Figure 239 shows a nucleotide sequence (SEQ ID NO:239) of a native sequence PRO1156 cDNA, wherein SEQ ID NO:239 is a clone designated herein as "DNA59853-1505".

Figure 240 shows the amino acid sequence (SEQ ID NO:240) derived from the coding sequence of SEQ ID NO:239 shown in Figure 239.

5 Figure 241 shows a nucleotide sequence (SEQ ID NO:241) of a native sequence PRO1098 cDNA, wherein SEQ ID NO:241 is a clone designated herein as "DNA59854-1459".

Figure 242 shows the amino acid sequence (SEQ ID NO:242) derived from the coding sequence of SEQ ID NO:241 shown in Figure 241.

Figure 243 shows a nucleotide sequence (SEQ ID NO:243) of a native sequence PRO1128 cDNA, wherein SEQ ID NO:243 is a clone designated herein as "DNA59855-1485".

10 Figure 244 shows the amino acid sequence (SEQ ID NO:244) derived from the coding sequence of SEQ ID NO:243 shown in Figure 243.

Figure 245 shows a nucleotide sequence (SEQ ID NO:245) of a native sequence PRO1248 cDNA, wherein SEQ ID NO:245 is a clone designated herein as "DNA60278-1530".

15 Figure 246 shows the amino acid sequence (SEQ ID NO:246) derived from the coding sequence of SEQ ID NO:245 shown in Figure 245.

Figure 247 shows a nucleotide sequence (SEQ ID NO:247) of a native sequence PRO1127 cDNA, wherein SEQ ID NO:247 is a clone designated herein as "DNA60283-1484".

Figure 248 shows the amino acid sequence (SEQ ID NO:248) derived from the coding sequence of SEQ ID NO:247 shown in Figure 247.

20 Figure 249 shows a nucleotide sequence (SEQ ID NO:249) of a native sequence PRO1316 cDNA, wherein SEQ ID NO:249 is a clone designated herein as "DNA60608-1577".

Figure 250 shows the amino acid sequence (SEQ ID NO:250) derived from the coding sequence of SEQ ID NO:249 shown in Figure 249.

25 Figure 251 shows a nucleotide sequence (SEQ ID NO:251) of a native sequence PRO1197 cDNA, wherein SEQ ID NO:251 is a clone designated herein as "DNA60611-1524".

Figure 252 shows the amino acid sequence (SEQ ID NO:252) derived from the coding sequence of SEQ ID NO:251 shown in Figure 251.

Figure 253 shows a nucleotide sequence (SEQ ID NO:253) of a native sequence PRO1125 cDNA, wherein SEQ ID NO:253 is a clone designated herein as "DNA60619-1482".

30 Figure 254 shows the amino acid sequence (SEQ ID NO:254) derived from the coding sequence of SEQ ID NO:253 shown in Figure 253.

Figure 255 shows a nucleotide sequence (SEQ ID NO:255) of a native sequence PRO1158 cDNA, wherein SEQ ID NO:255 is a clone designated herein as "DNA60625-1507".

35 Figure 256 shows the amino acid sequence (SEQ ID NO:256) derived from the coding sequence of SEQ ID NO:255 shown in Figure 255.

Figure 257 shows a nucleotide sequence (SEQ ID NO:257) of a native sequence PRO1124 cDNA, wherein SEQ ID NO:257 is a clone designated herein as "DNA60629-1481".

Figure 258 shows the amino acid sequence (SEQ ID NO:258) derived from the coding sequence of SEQ ID NO:257 shown in Figure 257.

Figure 259 shows a nucleotide sequence (SEQ ID NO:259) of a native sequence PRO1380 cDNA, wherein SEQ ID NO:259 is a clone designated herein as "DNA60740-1615".

5 Figure 260 shows the amino acid sequence (SEQ ID NO:260) derived from the coding sequence of SEQ ID NO:259 shown in Figure 259.

Figure 261 shows a nucleotide sequence (SEQ ID NO:261) of a native sequence PRO1377 cDNA, wherein SEQ ID NO:261 is a clone designated herein as "DNA61608-1606".

Figure 262 shows the amino acid sequence (SEQ ID NO:262) derived from the coding sequence of SEQ ID NO:261 shown in Figure 261.

10 Figure 263 shows a nucleotide sequence (SEQ ID NO:263) of a native sequence PRO1287 cDNA, wherein SEQ ID NO:263 is a clone designated herein as "DNA61755-1554".

Figure 264 shows the amino acid sequence (SEQ ID NO:264) derived from the coding sequence of SEQ ID NO:263 shown in Figure 263.

15 Figure 265 shows a nucleotide sequence (SEQ ID NO:265) of a native sequence PRO1249 cDNA, wherein SEQ ID NO:265 is a clone designated herein as "DNA62809-1531".

Figure 266 shows the amino acid sequence (SEQ ID NO:266) derived from the coding sequence of SEQ ID NO:265 shown in Figure 265.

Figure 267 shows a nucleotide sequence (SEQ ID NO:267) of a native sequence PRO1335 cDNA, wherein SEQ ID NO:267 is a clone designated herein as "DNA62812-1594".

20 Figure 268 shows the amino acid sequence (SEQ ID NO:268) derived from the coding sequence of SEQ ID NO:267 shown in Figure 267.

Figure 269 shows a nucleotide sequence (SEQ ID NO:269) of a native sequence PRO3572 cDNA, wherein SEQ ID NO:269 is a clone designated herein as "DNA62813-2544".

25 Figure 270 shows the amino acid sequence (SEQ ID NO:270) derived from the coding sequence of SEQ ID NO:269 shown in Figure 269.

Figure 271 shows a nucleotide sequence (SEQ ID NO:271) of a native sequence PRO1599 cDNA, wherein SEQ ID NO:271 is a clone designated herein as "DNA62845-1684".

Figure 272 shows the amino acid sequence (SEQ ID NO:272) derived from the coding sequence of SEQ ID NO:271 shown in Figure 271.

30 Figure 273 shows a nucleotide sequence (SEQ ID NO:273) of a native sequence PRO1374 cDNA, wherein SEQ ID NO:273 is a clone designated herein as "DNA64849-1604".

Figure 274 shows the amino acid sequence (SEQ ID NO:274) derived from the coding sequence of SEQ ID NO:273 shown in Figure 273.

35 Figure 275 shows a nucleotide sequence (SEQ ID NO:275) of a native sequence PRO1345 cDNA, wherein SEQ ID NO:275 is a clone designated herein as "DNA64852-1589".

Figure 276 shows the amino acid sequence (SEQ ID NO:276) derived from the coding sequence of SEQ ID NO:275 shown in Figure 275.

Figure 277 shows a nucleotide sequence (SEQ ID NO:277) of a native sequence PRO1311 cDNA, wherein SEQ ID NO:277 is a clone designated herein as "DNA64863-1573".

Figure 278 shows the amino acid sequence (SEQ ID NO:278) derived from the coding sequence of SEQ ID NO:277 shown in Figure 277.

5 Figure 279 shows a nucleotide sequence (SEQ ID NO:279) of a native sequence PRO1357 cDNA, wherein SEQ ID NO:279 is a clone designated herein as "DNA64881-1602".

Figure 280 shows the amino acid sequence (SEQ ID NO:280) derived from the coding sequence of SEQ ID NO:279 shown in Figure 279.

Figure 281 shows a nucleotide sequence (SEQ ID NO:281) of a native sequence PRO1557 cDNA, wherein SEQ ID NO:281 is a clone designated herein as "DNA64902-1667".

10 Figure 282 shows the amino acid sequence (SEQ ID NO:282) derived from the coding sequence of SEQ ID NO:281 shown in Figure 281.

Figure 283 shows a nucleotide sequence (SEQ ID NO:283) of a native sequence PRO1305 cDNA, wherein SEQ ID NO:283 is a clone designated herein as "DNA64952-1568".

15 Figure 284 shows the amino acid sequence (SEQ ID NO:284) derived from the coding sequence of SEQ ID NO:283 shown in Figure 283.

Figure 285 shows a nucleotide sequence (SEQ ID NO:285) of a native sequence PRO1302 cDNA, wherein SEQ ID NO:285 is a clone designated herein as "DNA65403-1565".

Figure 286 shows the amino acid sequence (SEQ ID NO:286) derived from the coding sequence of SEQ ID NO:285 shown in Figure 285.

20 Figure 287 shows a nucleotide sequence (SEQ ID NO:287) of a native sequence PRO1266 cDNA, wherein SEQ ID NO:287 is a clone designated herein as "DNA65413-1534".

Figure 288 shows the amino acid sequence (SEQ ID NO:288) derived from the coding sequence of SEQ ID NO:287 shown in Figure 287.

25 Figures 289A-289B show a nucleotide sequence (SEQ ID NO:289) of a native sequence PRO1336 cDNA, wherein SEQ ID NO:289 is a clone designated herein as "DNA65423-1595".

Figure 290 shows the amino acid sequence (SEQ ID NO:290) derived from the coding sequence of SEQ ID NO:289 shown in Figures 289A-289B.

Figure 291 shows a nucleotide sequence (SEQ ID NO:291) of a native sequence PRO1278 cDNA, wherein SEQ ID NO:291 is a clone designated herein as "DNA66304-1546".

30 Figure 292 shows the amino acid sequence (SEQ ID NO:292) derived from the coding sequence of SEQ ID NO:291 shown in Figure 291.

Figure 293 shows a nucleotide sequence (SEQ ID NO:293) of a native sequence PRO1270 cDNA, wherein SEQ ID NO:293 is a clone designated herein as "DNA66308-1537".

35 Figure 294 shows the amino acid sequence (SEQ ID NO:294) derived from the coding sequence of SEQ ID NO:293 shown in Figure 293.

Figure 295 shows a nucleotide sequence (SEQ ID NO:295) of a native sequence PRO1298 cDNA, wherein SEQ ID NO:295 is a clone designated herein as "DNA66511-1563".

Figure 296 shows the amino acid sequence (SEQ ID NO:296) derived from the coding sequence of SEQ ID NO:295 shown in Figure 295.

Figure 297 shows a nucleotide sequence (SEQ ID NO:297) of a native sequence PRO1301 cDNA, wherein SEQ ID NO:297 is a clone designated herein as "DNA66512-1564".

5 Figure 298 shows the amino acid sequence (SEQ ID NO:298) derived from the coding sequence of SEQ ID NO:297 shown in Figure 297.

Figure 299 shows a nucleotide sequence (SEQ ID NO:299) of a native sequence PRO1268 cDNA, wherein SEQ ID NO:299 is a clone designated herein as "DNA66519-1535".

Figure 300 shows the amino acid sequence (SEQ ID NO:300) derived from the coding sequence of SEQ ID NO:299 shown in Figure 299.

10 Figure 301 shows a nucleotide sequence (SEQ ID NO:301) of a native sequence PRO1327 cDNA, wherein SEQ ID NO:301 is a clone designated herein as "DNA66521-1583".

Figure 302 shows the amino acid sequence (SEQ ID NO:302) derived from the coding sequence of SEQ ID NO:301 shown in Figure 301.

15 Figure 303 shows a nucleotide sequence (SEQ ID NO:303) of a native sequence PRO1328 cDNA, wherein SEQ ID NO:303 is a clone designated herein as "DNA66658-1584".

Figure 304 shows the amino acid sequence (SEQ ID NO:304) derived from the coding sequence of SEQ ID NO:303 shown in Figure 303.

Figure 305 shows a nucleotide sequence (SEQ ID NO:305) of a native sequence PRO1329 cDNA, wherein SEQ ID NO:305 is a clone designated herein as "DNA66660-1585".

20 Figure 306 shows the amino acid sequence (SEQ ID NO:306) derived from the coding sequence of SEQ ID NO:305 shown in Figure 305.

Figure 307 shows a nucleotide sequence (SEQ ID NO:307) of a native sequence PRO1339 cDNA, wherein SEQ ID NO:307 is a clone designated herein as "DNA66669-1597".

25 Figure 308 shows the amino acid sequence (SEQ ID NO:308) derived from the coding sequence of SEQ ID NO:307 shown in Figure 307.

Figure 309 shows a nucleotide sequence (SEQ ID NO:309) of a native sequence PRO1342 cDNA, wherein SEQ ID NO:309 is a clone designated herein as "DNA66674-1599".

Figure 310 shows the amino acid sequence (SEQ ID NO:310) derived from the coding sequence of SEQ ID NO:309 shown in Figure 309.

30 Figures 311A-311B show a nucleotide sequence (SEQ ID NO:311) of a native sequence PRO1487 cDNA, wherein SEQ ID NO:311 is a clone designated herein as "DNA68836-1656".

Figure 312 shows the amino acid sequence (SEQ ID NO:312) derived from the coding sequence of SEQ ID NO:311 shown in Figures 311A-311B.

35 Figure 313 shows a nucleotide sequence (SEQ ID NO:313) of a native sequence PRO3579 cDNA, wherein SEQ ID NO:313 is a clone designated herein as "DNA68862-2546".

Figure 314 shows the amino acid sequence (SEQ ID NO:314) derived from the coding sequence of SEQ ID NO:313 shown in Figure 313.

Figure 315 shows a nucleotide sequence (SEQ ID NO:315) of a native sequence PRO1472 cDNA, wherein SEQ ID NO:315 is a clone designated herein as "DNA68866-1644".

Figure 316 shows the amino acid sequence (SEQ ID NO:316) derived from the coding sequence of SEQ ID NO:315 shown in Figure 315.

5 Figure 317 shows a nucleotide sequence (SEQ ID NO:317) of a native sequence PRO1385 cDNA, wherein SEQ ID NO:317 is a clone designated herein as "DNA68869-1610".

Figure 318 shows the amino acid sequence (SEQ ID NO:318) derived from the coding sequence of SEQ ID NO:317 shown in Figure 317.

Figure 319 shows a nucleotide sequence (SEQ ID NO:319) of a native sequence PRO1461 cDNA, wherein SEQ ID NO:319 is a clone designated herein as "DNA68871-1638".

10 Figure 320 shows the amino acid sequence (SEQ ID NO:320) derived from the coding sequence of SEQ ID NO:319 shown in Figure 319.

Figure 321 shows a nucleotide sequence (SEQ ID NO:321) of a native sequence PRO1429 cDNA, wherein SEQ ID NO:321 is a clone designated herein as "DNA68879-1631".

15 Figure 322 shows the amino acid sequence (SEQ ID NO:322) derived from the coding sequence of SEQ ID NO:321 shown in Figure 321.

Figure 323 shows a nucleotide sequence (SEQ ID NO:323) of a native sequence PRO1568 cDNA, wherein SEQ ID NO:323 is a clone designated herein as "DNA68880-1676".

Figure 324 shows the amino acid sequence (SEQ ID NO:324) derived from the coding sequence of SEQ ID NO:323 shown in Figure 323.

20 Figure 325 shows a nucleotide sequence (SEQ ID NO:325) of a native sequence PRO1569 cDNA, wherein SEQ ID NO:325 is a clone designated herein as "DNA68882-1677".

Figure 326 shows the amino acid sequence (SEQ ID NO:326) derived from the coding sequence of SEQ ID NO:325 shown in Figure 325.

25 Figure 327 shows a nucleotide sequence (SEQ ID NO:327) of a native sequence PRO1753 cDNA, wherein SEQ ID NO:327 is a clone designated herein as "DNA68883-1691".

Figure 328 shows the amino acid sequence (SEQ ID NO:328) derived from the coding sequence of SEQ ID NO:327 shown in Figure 327.

Figure 329 shows a nucleotide sequence (SEQ ID NO:329) of a native sequence PRO1570 cDNA, wherein SEQ ID NO:329 is a clone designated herein as "DNA68885-1678".

30 Figure 330 shows the amino acid sequence (SEQ ID NO:330) derived from the coding sequence of SEQ ID NO:329 shown in Figure 329.

Figure 331 shows a nucleotide sequence (SEQ ID NO:331) of a native sequence PRO1559 cDNA, wherein SEQ ID NO:331 is a clone designated herein as "DNA68886".

35 Figure 332 shows the amino acid sequence (SEQ ID NO:332) derived from the coding sequence of SEQ ID NO:331 shown in Figure 331.

Figure 333 shows a nucleotide sequence (SEQ ID NO:333) of a native sequence PRO1486 cDNA, wherein SEQ ID NO:333 is a clone designated herein as "DNA71180-1655".

Figure 334 shows the amino acid sequence (SEQ ID NO:334) derived from the coding sequence of SEQ ID NO:333 shown in Figure 333.

Figure 335 shows a nucleotide sequence (SEQ ID NO:335) of a native sequence PRO1433 cDNA, wherein SEQ ID NO:335 is a clone designated herein as "DNA71184-1634".

5 Figure 336 shows the amino acid sequence (SEQ ID NO:336) derived from the coding sequence of SEQ ID NO:335 shown in Figure 335.

Figure 337 shows a nucleotide sequence (SEQ ID NO:337) of a native sequence PRO1490 cDNA, wherein SEQ ID NO:337 is a clone designated herein as "DNA71213-1659".

Figure 338 shows the amino acid sequence (SEQ ID NO:338) derived from the coding sequence of SEQ ID NO:337 shown in Figure 337.

10 Figure 339 shows a nucleotide sequence (SEQ ID NO:339) of a native sequence PRO1482 cDNA, wherein SEQ ID NO:339 is a clone designated herein as "DNA71234-1651".

Figure 340 shows the amino acid sequence (SEQ ID NO:340) derived from the coding sequence of SEQ ID NO:339 shown in Figure 339.

15 Figure 341 shows a nucleotide sequence (SEQ ID NO:341) of a native sequence PRO1409 cDNA, wherein SEQ ID NO:341 is a clone designated herein as "DNA71269-1621".

Figure 342 shows the amino acid sequence (SEQ ID NO:342) derived from the coding sequence of SEQ ID NO:341 shown in Figure 341.

Figure 343 shows a nucleotide sequence (SEQ ID NO:343) of a native sequence PRO1446 cDNA, wherein SEQ ID NO:343 is a clone designated herein as "DNA71277-1636".

20 Figure 344 shows the amino acid sequence (SEQ ID NO:344) derived from the coding sequence of SEQ ID NO:343 shown in Figure 343.

Figure 345 shows a nucleotide sequence (SEQ ID NO:345) of a native sequence PRO1604 cDNA, wherein SEQ ID NO:345 is a clone designated herein as "DNA71286-1687".

25 Figure 346 shows the amino acid sequence (SEQ ID NO:346) derived from the coding sequence of SEQ ID NO:345 shown in Figure 345.

Figure 347 shows a nucleotide sequence (SEQ ID NO:347) of a native sequence PRO1491 cDNA, wherein SEQ ID NO:347 is a clone designated herein as "DNA71883-1660".

Figure 348 shows the amino acid sequence (SEQ ID NO:348) derived from the coding sequence of SEQ ID NO:347 shown in Figure 347.

30 Figure 349 shows a nucleotide sequence (SEQ ID NO:349) of a native sequence PRO1431 cDNA, wherein SEQ ID NO:349 is a clone designated herein as "DNA73401-1633".

Figure 350 shows the amino acid sequence (SEQ ID NO:350) derived from the coding sequence of SEQ ID NO:349 shown in Figure 349.

35 Figures 351A-351B show a nucleotide sequence (SEQ ID NO:351) of a native sequence PRO1563 cDNA, wherein SEQ ID NO:351 is a clone designated herein as "DNA73492-1671".

Figure 352 shows the amino acid sequence (SEQ ID NO:352) derived from the coding sequence of SEQ ID NO:351 shown in Figures 351A-351B.

Figure 353 shows a nucleotide sequence (SEQ ID NO:353) of a native sequence PRO1571 cDNA, wherein SEQ ID NO:353 is a clone designated herein as "DNA73730-1679".

Figure 354 shows the amino acid sequence (SEQ ID NO:354) derived from the coding sequence of SEQ ID NO:353 shown in Figure 353.

5 Figure 355 shows a nucleotide sequence (SEQ ID NO:355) of a native sequence PRO1572 cDNA, wherein SEQ ID NO:355 is a clone designated herein as "DNA73734-1680".

Figure 356 shows the amino acid sequence (SEQ ID NO:356) derived from the coding sequence of SEQ ID NO:355 shown in Figure 355.

Figure 357 shows a nucleotide sequence (SEQ ID NO:357) of a native sequence PRO1573 cDNA, wherein SEQ ID NO:357 is a clone designated herein as "DNA73735-1681".

10 Figure 358 shows the amino acid sequence (SEQ ID NO:358) derived from the coding sequence of SEQ ID NO:357 shown in Figure 357.

Figure 359 shows a nucleotide sequence (SEQ ID NO:359) of a native sequence PRO1508 cDNA, wherein SEQ ID NO:359 is a clone designated herein as "DNA73742-1662".

15 Figure 360 shows the amino acid sequence (SEQ ID NO:360) derived from the coding sequence of SEQ ID NO:359 shown in Figure 359.

Figure 361 shows a nucleotide sequence (SEQ ID NO:361) of a native sequence PRO1485 cDNA, wherein SEQ ID NO:361 is a clone designated herein as "DNA73746-1654".

Figure 362 shows the amino acid sequence (SEQ ID NO:362) derived from the coding sequence of SEQ ID NO:361 shown in Figure 361.

20 Figure 363 shows a nucleotide sequence (SEQ ID NO:363) of a native sequence PRO1564 cDNA, wherein SEQ ID NO:363 is a clone designated herein as "DNA73760-1672".

Figure 364 shows the amino acid sequence (SEQ ID NO:364) derived from the coding sequence of SEQ ID NO:363 shown in Figure 363.

25 Figure 365 shows a nucleotide sequence (SEQ ID NO:365) of a native sequence PRO1550 cDNA, wherein SEQ ID NO:365 is a clone designated herein as "DNA76393-1664".

Figure 366 shows the amino acid sequence (SEQ ID NO:366) derived from the coding sequence of SEQ ID NO:365 shown in Figure 365.

Figure 367 shows a nucleotide sequence (SEQ ID NO:367) of a native sequence PRO1757 cDNA, wherein SEQ ID NO:367 is a clone designated herein as "DNA76398-1699".

30 Figure 368 shows the amino acid sequence (SEQ ID NO:368) derived from the coding sequence of SEQ ID NO:367 shown in Figure 367.

Figure 369 shows a nucleotide sequence (SEQ ID NO:369) of a native sequence PRO1758 cDNA, wherein SEQ ID NO:369 is a clone designated herein as "DNA76399-1700".

35 Figure 370 shows the amino acid sequence (SEQ ID NO:370) derived from the coding sequence of SEQ ID NO:369 shown in Figure 369.

Figure 371 shows a nucleotide sequence (SEQ ID NO:371) of a native sequence PRO1781 cDNA, wherein SEQ ID NO:371 is a clone designated herein as "DNA76522-2500".

Figure 372 shows the amino acid sequence (SEQ ID NO:372) derived from the coding sequence of SEQ ID NO:371 shown in Figure 371.

Figure 373 shows a nucleotide sequence (SEQ ID NO:373) of a native sequence PRO1606 cDNA, wherein SEQ ID NO:373 is a clone designated herein as "DNA76533-1689".

5 Figure 374 shows the amino acid sequence (SEQ ID NO:374) derived from the coding sequence of SEQ ID NO:373 shown in Figure 373.

Figure 375 shows a nucleotide sequence (SEQ ID NO:375) of a native sequence PRO1784 cDNA, wherein SEQ ID NO:375 is a clone designated herein as "DNA77303-2502".

Figure 376 shows the amino acid sequence (SEQ ID NO:376) derived from the coding sequence of SEQ ID NO:375 shown in Figure 375.

10 Figure 377 shows a nucleotide sequence (SEQ ID NO:377) of a native sequence PRO1774 cDNA, wherein SEQ ID NO:377 is a clone designated herein as "DNA77626-1705".

Figure 378 shows the amino acid sequence (SEQ ID NO:378) derived from the coding sequence of SEQ ID NO:377 shown in Figure 377.

15 Figure 379 shows a nucleotide sequence (SEQ ID NO:379) of a native sequence PRO1605 cDNA, wherein SEQ ID NO:379 is a clone designated herein as "DNA77648-1688".

Figure 380 shows the amino acid sequence (SEQ ID NO:380) derived from the coding sequence of SEQ ID NO:379 shown in Figure 379.

Figure 381 shows a nucleotide sequence (SEQ ID NO:381) of a native sequence PRO1928 cDNA, wherein SEQ ID NO:381 is a clone designated herein as "DNA81754-2532".

20 Figure 382 shows the amino acid sequence (SEQ ID NO:382) derived from the coding sequence of SEQ ID NO:381 shown in Figure 381.

Figure 383 shows a nucleotide sequence (SEQ ID NO:383) of a native sequence PRO1865 cDNA, wherein SEQ ID NO:383 is a clone designated herein as "DNA81757-2512".

25 Figure 384 shows the amino acid sequence (SEQ ID NO:384) derived from the coding sequence of SEQ ID NO:383 shown in Figure 383.

Figure 385 shows a nucleotide sequence (SEQ ID NO:385) of a native sequence PRO1925 cDNA, wherein SEQ ID NO:385 is a clone designated herein as "DNA82302-2529".

Figure 386 shows the amino acid sequence (SEQ ID NO:386) derived from the coding sequence of SEQ ID NO:385 shown in Figure 385.

30 Figure 387 shows a nucleotide sequence (SEQ ID NO:387) of a native sequence PRO1926 cDNA, wherein SEQ ID NO:387 is a clone designated herein as "DNA82340-2530".

Figure 388 shows the amino acid sequence (SEQ ID NO:388) derived from the coding sequence of SEQ ID NO:387 shown in Figure 387.

35 Figure 389 shows a nucleotide sequence (SEQ ID NO:389) of a native sequence PRO2630 cDNA, wherein SEQ ID NO:389 is a clone designated herein as "DNA83551".

Figure 390 shows the amino acid sequence (SEQ ID NO:390) derived from the coding sequence of SEQ ID NO:389 shown in Figure 389.



Figure 391 shows a nucleotide sequence (SEQ ID NO:391) of a native sequence PRO3443 cDNA, wherein SEQ ID NO:391 is a clone designated herein as "DNA87991-2540".

Figure 392 shows the amino acid sequence (SEQ ID NO:392) derived from the coding sequence of SEQ ID NO:391 shown in Figure 391.

5 Figure 393 shows a nucleotide sequence (SEQ ID NO:393) of a native sequence PRO3301 cDNA, wherein SEQ ID NO:393 is a clone designated herein as "DNA88002".

Figure 394 shows the amino acid sequence (SEQ ID NO:394) derived from the coding sequence of SEQ ID NO:393 shown in Figure 393.

Figure 395 shows a nucleotide sequence (SEQ ID NO:395) of a native sequence PRO3442 cDNA, wherein SEQ ID NO:395 is a clone designated herein as "DNA92238-2539".

10 Figure 396 shows the amino acid sequence (SEQ ID NO:396) derived from the coding sequence of SEQ ID NO:395 shown in Figure 395.

Figure 397 shows a nucleotide sequence (SEQ ID NO:397) of a native sequence PRO4978 cDNA, wherein SEQ ID NO:397 is a clone designated herein as "DNA95930".

15 Figure 398 shows the amino acid sequence (SEQ ID NO:398) derived from the coding sequence of SEQ ID NO:397 shown in Figure 397.

Figure 399 shows a nucleotide sequence (SEQ ID NO:399) of a native sequence PRO5801 cDNA, wherein SEQ ID NO:399 is a clone designated herein as "DNA115291-2681".

Figure 400 shows the amino acid sequence (SEQ ID NO:400) derived from the coding sequence of SEQ ID NO:399 shown in Figure 399.

20 Figure 401 shows a nucleotide sequence (SEQ ID NO:401) of a native sequence PRO19630 cDNA, wherein SEQ ID NO:401 is a clone designated herein as "DNA23336-2861".

Figure 402 shows the amino acid sequence (SEQ ID NO:402) derived from the coding sequence of SEQ ID NO:401 shown in Figure 401.

25 Figure 403 shows a nucleotide sequence (SEQ ID NO:403) of a native sequence PRO203 cDNA, wherein SEQ ID NO:403 is a clone designated herein as "DNA30862-1396".

Figure 404 shows the amino acid sequence (SEQ ID NO:404) derived from the coding sequence of SEQ ID NO:403 shown in Figure 403.

Figure 405 shows a nucleotide sequence (SEQ ID NO:405) of a native sequence PRO204 cDNA, wherein SEQ ID NO:405 is a clone designated herein as "DNA30871-1157".

30 Figure 406 shows the amino acid sequence (SEQ ID NO:406) derived from the coding sequence of SEQ ID NO:405 shown in Figure 405.

Figure 407 shows a nucleotide sequence (SEQ ID NO:407) of a native sequence PRO210 cDNA, wherein SEQ ID NO:407 is a clone designated herein as "DNA32279-1131".

35 Figure 408 shows the amino acid sequence (SEQ ID NO:408) derived from the coding sequence of SEQ ID NO:407 shown in Figure 407.

Figure 409 shows a nucleotide sequence (SEQ ID NO:409) of a native sequence PRO223 cDNA, wherein SEQ ID NO:409 is a clone designated herein as "DNA33206-1165".

Figure 410 shows the amino acid sequence (SEQ ID NO:410) derived from the coding sequence of SEQ ID NO:409 shown in Figure 409.

Figure 411 shows a nucleotide sequence (SEQ ID NO:411) of a native sequence PRO247 cDNA, wherein SEQ ID NO:411 is a clone designated herein as "DNA35673-1201".

5      Figure 412 shows the amino acid sequence (SEQ ID NO:412) derived from the coding sequence of SEQ ID NO:411 shown in Figure 411.

Figure 413 shows a nucleotide sequence (SEQ ID NO:413) of a native sequence PRO358 cDNA, wherein SEQ ID NO:413 is a clone designated herein as "DNA47361-1154-2".

Figure 414 shows the amino acid sequence (SEQ ID NO:414) derived from the coding sequence of SEQ ID NO:413 shown in Figure 413.

10      Figure 415 shows a nucleotide sequence (SEQ ID NO:415) of a native sequence PRO724 cDNA, wherein SEQ ID NO:415 is a clone designated herein as "DNA49631-1328".

Figure 416 shows the amino acid sequence (SEQ ID NO:416) derived from the coding sequence of SEQ ID NO:415 shown in Figure 415.

15      Figure 417 shows a nucleotide sequence (SEQ ID NO:417) of a native sequence PRO868 cDNA, wherein SEQ ID NO:417 is a clone designated herein as "DNA52594-1270".

Figure 418 shows the amino acid sequence (SEQ ID NO:418) derived from the coding sequence of SEQ ID NO:417 shown in Figure 417.

Figure 419 shows a nucleotide sequence (SEQ ID NO:419) of a native sequence PRO740 cDNA, wherein SEQ ID NO:419 is a clone designated herein as "DNA55800-1263".

20      Figure 420 shows the amino acid sequence (SEQ ID NO:420) derived from the coding sequence of SEQ ID NO:419 shown in Figure 419.

Figure 421 shows a nucleotide sequence (SEQ ID NO:421) of a native sequence PRO1478 cDNA, wherein SEQ ID NO:421 is a clone designated herein as "DNA56531-1648".

25      Figure 422 shows the amino acid sequence (SEQ ID NO:422) derived from the coding sequence of SEQ ID NO:421 shown in Figure 421.

Figure 423 shows a nucleotide sequence (SEQ ID NO:423) of a native sequence PRO162 cDNA, wherein SEQ ID NO:423 is a clone designated herein as "DNA56965-1356".

Figure 424 shows the amino acid sequence (SEQ ID NO:424) derived from the coding sequence of SEQ ID NO:423 shown in Figure 423.

30      Figure 425 shows a nucleotide sequence (SEQ ID NO:425) of a native sequence PRO828 cDNA, wherein SEQ ID NO:425 is a clone designated herein as "DNA57037-1444".

Figure 426 shows the amino acid sequence (SEQ ID NO:426) derived from the coding sequence of SEQ ID NO:425 shown in Figure 425.

35      Figure 427 shows a nucleotide sequence (SEQ ID NO:427) of a native sequence PRO819 cDNA, wherein SEQ ID NO:427 is a clone designated herein as "DNA57695-1340".

Figure 428 shows the amino acid sequence (SEQ ID NO:428) derived from the coding sequence of SEQ ID NO:427 shown in Figure 427.

Figure 429 shows a nucleotide sequence (SEQ ID NO:429) of a native sequence PRO813 cDNA, wherein SEQ ID NO:429 is a clone designated herein as "DNA57834-1339".

Figure 430 shows the amino acid sequence (SEQ ID NO:430) derived from the coding sequence of SEQ ID NO:429 shown in Figure 429.

5 Figure 431 shows a nucleotide sequence (SEQ ID NO:431) of a native sequence PRO1194 cDNA, wherein SEQ ID NO:431 is a clone designated herein as "DNA57841-1522".

Figure 432 shows the amino acid sequence (SEQ ID NO:432) derived from the coding sequence of SEQ ID NO:431 shown in Figure 431.

Figure 433 shows a nucleotide sequence (SEQ ID NO:433) of a native sequence PRO887 cDNA, wherein SEQ ID NO:433 is a clone designated herein as "DNA58130".

10 Figure 434 shows the amino acid sequence (SEQ ID NO:434) derived from the coding sequence of SEQ ID NO:433 shown in Figure 433.

Figure 435 shows a nucleotide sequence (SEQ ID NO:435) of a native sequence PRO1071 cDNA, wherein SEQ ID NO:435 is a clone designated herein as "DNA58847-1383".

15 Figure 436 shows the amino acid sequence (SEQ ID NO:436) derived from the coding sequence of SEQ ID NO:435 shown in Figure 435.

Figure 437 shows a nucleotide sequence (SEQ ID NO:437) of a native sequence PRO1029 cDNA, wherein SEQ ID NO:437 is a clone designated herein as "DNA59493-1420".

Figure 438 shows the amino acid sequence (SEQ ID NO:438) derived from the coding sequence of SEQ ID NO:437 shown in Figure 437.

20 Figure 439 shows a nucleotide sequence (SEQ ID NO:439) of a native sequence PRO1190 cDNA, wherein SEQ ID NO:439 is a clone designated herein as "DNA59586-1520".

Figure 440 shows the amino acid sequence (SEQ ID NO:440) derived from the coding sequence of SEQ ID NO:439 shown in Figure 439.

25 Figure 441 shows a nucleotide sequence (SEQ ID NO:441) of a native sequence PRO4334 cDNA, wherein SEQ ID NO:441 is a clone designated herein as "DNA59608-2577".

Figure 442 shows the amino acid sequence (SEQ ID NO:442) derived from the coding sequence of SEQ ID NO:441 shown in Figure 441.

Figure 443 shows a nucleotide sequence (SEQ ID NO:443) of a native sequence PRO1155 cDNA, wherein SEQ ID NO:443 is a clone designated herein as "DNA59849-1504".

30 Figure 444 shows the amino acid sequence (SEQ ID NO:444) derived from the coding sequence of SEQ ID NO:443 shown in Figure 443.

Figure 445 shows a nucleotide sequence (SEQ ID NO:445) of a native sequence PRO1157 cDNA, wherein SEQ ID NO:445 is a clone designated herein as "DNA60292-1506".

35 Figure 446 shows the amino acid sequence (SEQ ID NO:446) derived from the coding sequence of SEQ ID NO:445 shown in Figure 445.

Figure 447 shows a nucleotide sequence (SEQ ID NO:447) of a native sequence PRO1122 cDNA, wherein SEQ ID NO:447 is a clone designated herein as "DNA62377-1381-1".

Figure 448 shows the amino acid sequence (SEQ ID NO:448) derived from the coding sequence of SEQ ID NO:447 shown in Figure 447.

Figure 449 shows a nucleotide sequence (SEQ ID NO:449) of a native sequence PRO1183 cDNA, wherein SEQ ID NO:449 is a clone designated herein as "DNA62880-1513".

5 Figure 450 shows the amino acid sequence (SEQ ID NO:450) derived from the coding sequence of SEQ ID NO:449 shown in Figure 449.

Figure 451 shows a nucleotide sequence (SEQ ID NO:451) of a native sequence PRO1337 cDNA, wherein SEQ ID NO:451 is a clone designated herein as "DNA66672-1586".

Figure 452 shows the amino acid sequence (SEQ ID NO:452) derived from the coding sequence of SEQ ID NO:451 shown in Figure 451.

10 Figure 453 shows a nucleotide sequence (SEQ ID NO:453) of a native sequence PRO1480 cDNA, wherein SEQ ID NO:453 is a clone designated herein as "DNA67962-1649".

Figure 454 shows the amino acid sequence (SEQ ID NO:454) derived from the coding sequence of SEQ ID NO:453 shown in Figure 453.

15 Figure 455 shows a nucleotide sequence (SEQ ID NO:455) of a native sequence PRO19645 cDNA, wherein SEQ ID NO:455 is a clone designated herein as "DNA69555-2867".

Figure 456 shows the amino acid sequence (SEQ ID NO:456) derived from the coding sequence of SEQ ID NO:455 shown in Figure 455.

Figure 457 shows a nucleotide sequence (SEQ ID NO:457) of a native sequence PRO9782 cDNA, wherein SEQ ID NO:457 is a clone designated herein as "DNA71162-2764".

20 Figure 458 shows the amino acid sequence (SEQ ID NO:458) derived from the coding sequence of SEQ ID NO:457 shown in Figure 457.

Figure 459 shows a nucleotide sequence (SEQ ID NO:459) of a native sequence PRO1419 cDNA, wherein SEQ ID NO:459 is a clone designated herein as "DNA71290-1630".

25 Figure 460 shows the amino acid sequence (SEQ ID NO:460) derived from the coding sequence of SEQ ID NO:459 shown in Figure 459.

Figure 461 shows a nucleotide sequence (SEQ ID NO:461) of a native sequence PRO1575 cDNA, wherein SEQ ID NO:461 is a clone designated herein as "DNA76401-1683".

Figure 462 shows the amino acid sequence (SEQ ID NO:462) derived from the coding sequence of SEQ ID NO:461 shown in Figure 461.

30 Figure 463 shows a nucleotide sequence (SEQ ID NO:463) of a native sequence PRO1567 cDNA, wherein SEQ ID NO:463 is a clone designated herein as "DNA76541-1675".

Figure 464 shows the amino acid sequence (SEQ ID NO:464) derived from the coding sequence of SEQ ID NO:463 shown in Figure 463.

35 Figure 465 shows a nucleotide sequence (SEQ ID NO:465) of a native sequence PRO1891 cDNA, wherein SEQ ID NO:465 is a clone designated herein as "DNA76788-2526".

Figure 466 shows the amino acid sequence (SEQ ID NO:466) derived from the coding sequence of SEQ ID NO:465 shown in Figure 465.

Figure 467 shows a nucleotide sequence (SEQ ID NO:467) of a native sequence PRO1889 cDNA, wherein SEQ ID NO:467 is a clone designated herein as "DNA77623-2524".

Figure 468 shows the amino acid sequence (SEQ ID NO:468) derived from the coding sequence of SEQ ID NO:467 shown in Figure 467.

5 Figure 469 shows a nucleotide sequence (SEQ ID NO:469) of a native sequence PRO1785 cDNA, wherein SEQ ID NO:469 is a clone designated herein as "DNA80136-2503".

Figure 470 shows the amino acid sequence (SEQ ID NO:470) derived from the coding sequence of SEQ ID NO:469 shown in Figure 469.

Figure 471 shows a nucleotide sequence (SEQ ID NO:471) of a native sequence PRO6003 cDNA, wherein SEQ ID NO:471 is a clone designated herein as "DNA83568-2692".

10 Figure 472 shows the amino acid sequence (SEQ ID NO:472) derived from the coding sequence of SEQ ID NO:471 shown in Figure 471.

Figure 473 shows a nucleotide sequence (SEQ ID NO:473) of a native sequence PRO4333 cDNA, wherein SEQ ID NO:473 is a clone designated herein as "DNA84210-2576".

15 Figure 474 shows the amino acid sequence (SEQ ID NO:474) derived from the coding sequence of SEQ ID NO:473 shown in Figure 473.

Figure 475 shows a nucleotide sequence (SEQ ID NO:475) of a native sequence PRO4356 cDNA, wherein SEQ ID NO:475 is a clone designated herein as "DNA86576-2595".

Figure 476 shows the amino acid sequence (SEQ ID NO:476) derived from the coding sequence of SEQ ID NO:475 shown in Figure 475.

20 Figure 477 shows a nucleotide sequence (SEQ ID NO:477) of a native sequence PRO4352 cDNA, wherein SEQ ID NO:477 is a clone designated herein as "DNA87976-2593".

Figure 478 shows the amino acid sequence (SEQ ID NO:478) derived from the coding sequence of SEQ ID NO:477 shown in Figure 477.

25 Figure 479 shows a nucleotide sequence (SEQ ID NO:479) of a native sequence PRO4354 cDNA, wherein SEQ ID NO:479 is a clone designated herein as "DNA92256-2596".

Figure 480 shows the amino acid sequence (SEQ ID NO:480) derived from the coding sequence of SEQ ID NO:479 shown in Figure 479.

Figure 481 shows a nucleotide sequence (SEQ ID NO:481) of a native sequence PRO4369 cDNA, wherein SEQ ID NO:481 is a clone designated herein as "DNA92289-2598".

30 Figure 482 shows the amino acid sequence (SEQ ID NO:482) derived from the coding sequence of SEQ ID NO:481 shown in Figure 481.

Figure 483 shows a nucleotide sequence (SEQ ID NO:483) of a native sequence PRO6030 cDNA, wherein SEQ ID NO:483 is a clone designated herein as "DNA96850-2705".

35 Figure 484 shows the amino acid sequence (SEQ ID NO:484) derived from the coding sequence of SEQ ID NO:483 shown in Figure 483.

Figure 485 shows a nucleotide sequence (SEQ ID NO:485) of a native sequence PRO4433 cDNA, wherein SEQ ID NO:485 is a clone designated herein as "DNA96855-2629".

Figure 486 shows the amino acid sequence (SEQ ID NO:486) derived from the coding sequence of SEQ ID NO:485 shown in Figure 485.

Figure 487 shows a nucleotide sequence (SEQ ID NO:487) of a native sequence PRO4424 cDNA, wherein SEQ ID NO:487 is a clone designated herein as "DNA96857-2636".

5 Figure 488 shows the amino acid sequence (SEQ ID NO:488) derived from the coding sequence of SEQ ID NO:487 shown in Figure 487.

Figure 489 shows a nucleotide sequence (SEQ ID NO:489) of a native sequence PRO6017 cDNA, wherein SEQ ID NO:489 is a clone designated herein as "DNA96860-2700".

Figure 490 shows the amino acid sequence (SEQ ID NO:490) derived from the coding sequence of SEQ ID NO:489 shown in Figure 489.

10 Figure 491 shows a nucleotide sequence (SEQ ID NO:491) of a native sequence PRO19563 cDNA, wherein SEQ ID NO:491 is a clone designated herein as "DNA96861-2844".

Figure 492 shows the amino acid sequence (SEQ ID NO:492) derived from the coding sequence of SEQ ID NO:491 shown in Figure 491.

15 Figure 493 shows a nucleotide sequence (SEQ ID NO:493) of a native sequence PRO6015 cDNA, wherein SEQ ID NO:493 is a clone designated herein as "DNA96866-2698".

Figure 494 shows the amino acid sequence (SEQ ID NO:494) derived from the coding sequence of SEQ ID NO:493 shown in Figure 493.

Figure 495 shows a nucleotide sequence (SEQ ID NO:495) of a native sequence PRO5779 cDNA, wherein SEQ ID NO:495 is a clone designated herein as "DNA96870-2676".

20 Figure 496 shows the amino acid sequence (SEQ ID NO:496) derived from the coding sequence of SEQ ID NO:495 shown in Figure 495.

Figure 497 shows a nucleotide sequence (SEQ ID NO:497) of a native sequence PRO5776 cDNA, wherein SEQ ID NO:497 is a clone designated herein as "DNA96872-2674".

25 Figure 498 shows the amino acid sequence (SEQ ID NO:498) derived from the coding sequence of SEQ ID NO:497 shown in Figure 497.

Figure 499 shows a nucleotide sequence (SEQ ID NO:499) of a native sequence PRO4430 cDNA, wherein SEQ ID NO:499 is a clone designated herein as "DNA96878-2626".

Figure 500 shows the amino acid sequence (SEQ ID NO:500) derived from the coding sequence of SEQ ID NO:499 shown in Figure 499.

30 Figure 501 shows a nucleotide sequence (SEQ ID NO:501) of a native sequence PRO4421 cDNA, wherein SEQ ID NO:501 is a clone designated herein as "DNA96879-2619".

Figure 502 shows the amino acid sequence (SEQ ID NO:502) derived from the coding sequence of SEQ ID NO:501 shown in Figure 501.

35 Figure 503 shows a nucleotide sequence (SEQ ID NO:503) of a native sequence PRO4499 cDNA, wherein SEQ ID NO:503 is a clone designated herein as "DNA96889-2641".

Figure 504 shows the amino acid sequence (SEQ ID NO:504) derived from the coding sequence of SEQ ID NO:503 shown in Figure 503.

Figure 505 shows a nucleotide sequence (SEQ ID NO:505) of a native sequence PRO4423 cDNA, wherein SEQ ID NO:505 is a clone designated herein as "DNA96893-2621".

Figure 506 shows the amino acid sequence (SEQ ID NO:506) derived from the coding sequence of SEQ ID NO:505 shown in Figure 505.

5 Figure 507 shows a nucleotide sequence (SEQ ID NO:507) of a native sequence PRO5998 cDNA, wherein SEQ ID NO:507 is a clone designated herein as "DNA96897-2688".

Figure 508 shows the amino acid sequence (SEQ ID NO:508) derived from the coding sequence of SEQ ID NO:507 shown in Figure 507.

Figure 509 shows a nucleotide sequence (SEQ ID NO:509) of a native sequence PRO4501 cDNA, wherein SEQ ID NO:509 is a clone designated herein as "DNA98564-2643".

10 Figure 510 shows the amino acid sequence (SEQ ID NO:510) derived from the coding sequence of SEQ ID NO:509 shown in Figure 509.

Figure 511 shows a nucleotide sequence (SEQ ID NO:511) of a native sequence PRO6240 cDNA, wherein SEQ ID NO:511 is a clone designated herein as "DNA107443-2718".

15 Figure 512 shows the amino acid sequence (SEQ ID NO:512) derived from the coding sequence of SEQ ID NO:511 shown in Figure 511.

Figure 513 shows a nucleotide sequence (SEQ ID NO:513) of a native sequence PRO6245 cDNA, wherein SEQ ID NO:513 is a clone designated herein as "DNA107786-2723".

Figure 514 shows the amino acid sequence (SEQ ID NO:514) derived from the coding sequence of SEQ ID NO:513 shown in Figure 513.

20 Figure 515 shows a nucleotide sequence (SEQ ID NO:515) of a native sequence PRO6175 cDNA, wherein SEQ ID NO:515 is a clone designated herein as "DNA108682-2712".

Figure 516 shows the amino acid sequence (SEQ ID NO:516) derived from the coding sequence of SEQ ID NO:515 shown in Figure 515.

25 Figure 517 shows a nucleotide sequence (SEQ ID NO:517) of a native sequence PRO9742 cDNA, wherein SEQ ID NO:517 is a clone designated herein as "DNA108684-2761".

Figure 518 shows the amino acid sequence (SEQ ID NO:518) derived from the coding sequence of SEQ ID NO:517 shown in Figure 517.

Figure 519 shows a nucleotide sequence (SEQ ID NO:519) of a native sequence PRO7179 cDNA, wherein SEQ ID NO:519 is a clone designated herein as "DNA108701-2749".

30 Figure 520 shows the amino acid sequence (SEQ ID NO:520) derived from the coding sequence of SEQ ID NO:519 shown in Figure 519.

Figure 521 shows a nucleotide sequence (SEQ ID NO:521) of a native sequence PRO6239 cDNA, wherein SEQ ID NO:521 is a clone designated herein as "DNA108720-2717".

35 Figure 522 shows the amino acid sequence (SEQ ID NO:522) derived from the coding sequence of SEQ ID NO:521 shown in Figure 521.

Figure 523 shows a nucleotide sequence (SEQ ID NO:523) of a native sequence PRO6493 cDNA, wherein SEQ ID NO:523 is a clone designated herein as "DNA108726-2729".

Figure 524 shows the amino acid sequence (SEQ ID NO:524) derived from the coding sequence of SEQ ID NO:523 shown in Figure 523.

Figures 525A-525B show a nucleotide sequence (SEQ ID NO:525) of a native sequence PRO9741 cDNA, wherein SEQ ID NO:525 is a clone designated herein as "DNA108728-2760".

5 Figure 526 shows the amino acid sequence (SEQ ID NO:526) derived from the coding sequence of SEQ ID NO:525 shown in Figures 525A-525B.

Figure 527 shows a nucleotide sequence (SEQ ID NO:527) of a native sequence PRO9822 cDNA, wherein SEQ ID NO:527 is a clone designated herein as "DNA108738-2767".

Figure 528 shows the amino acid sequence (SEQ ID NO:528) derived from the coding sequence of SEQ ID NO:527 shown in Figure 527.

10 Figure 529 shows a nucleotide sequence (SEQ ID NO:529) of a native sequence PRO6244 cDNA, wherein SEQ ID NO:529 is a clone designated herein as "DNA108743-2722".

Figure 530 shows the amino acid sequence (SEQ ID NO:530) derived from the coding sequence of SEQ ID NO:529 shown in Figure 529.

15 Figure 531 shows a nucleotide sequence (SEQ ID NO:531) of a native sequence PRO9740 cDNA, wherein SEQ ID NO:531 is a clone designated herein as "DNA108758-2759".

Figure 532 shows the amino acid sequence (SEQ ID NO:532) derived from the coding sequence of SEQ ID NO:531 shown in Figure 531.

Figure 533 shows a nucleotide sequence (SEQ ID NO:533) of a native sequence PRO9739 cDNA, wherein SEQ ID NO:533 is a clone designated herein as "DNA108765-2758".

20 Figure 534 shows the amino acid sequence (SEQ ID NO:534) derived from the coding sequence of SEQ ID NO:533 shown in Figure 533.

Figure 535 shows a nucleotide sequence (SEQ ID NO:535) of a native sequence PRO7177 cDNA, wherein SEQ ID NO:535 is a clone designated herein as "DNA108783-2747".

25 Figure 536 shows the amino acid sequence (SEQ ID NO:536) derived from the coding sequence of SEQ ID NO:535 shown in Figure 535.

Figure 537 shows a nucleotide sequence (SEQ ID NO:537) of a native sequence PRO7178 cDNA, wherein SEQ ID NO:537 is a clone designated herein as "DNA108789-2748".

Figure 538 shows the amino acid sequence (SEQ ID NO:538) derived from the coding sequence of SEQ ID NO:537 shown in Figure 537.

30 Figure 539 shows a nucleotide sequence (SEQ ID NO:539) of a native sequence PRO6246 cDNA, wherein SEQ ID NO:539 is a clone designated herein as "DNA108806-2724".

Figure 540 shows the amino acid sequence (SEQ ID NO:540) derived from the coding sequence of SEQ ID NO:539 shown in Figure 539.

35 Figure 541 shows a nucleotide sequence (SEQ ID NO:541) of a native sequence PRO6241 cDNA, wherein SEQ ID NO:541 is a clone designated herein as "DNA108936-2719".

Figure 542 shows the amino acid sequence (SEQ ID NO:542) derived from the coding sequence of SEQ ID NO:541 shown in Figure 541.



Figure 543 shows a nucleotide sequence (SEQ ID NO:543) of a native sequence PRO9835 cDNA, wherein SEQ ID NO:543 is a clone designated herein as "DNA119510-2771".

Figure 544 shows the amino acid sequence (SEQ ID NO:544) derived from the coding sequence of SEQ ID NO:543 shown in Figure 543.

5 Figure 545 shows a nucleotide sequence (SEQ ID NO:545) of a native sequence PRO9857 cDNA, wherein SEQ ID NO:545 is a clone designated herein as "DNA119517-2778".

Figure 546 shows the amino acid sequence (SEQ ID NO:546) derived from the coding sequence of SEQ ID NO:545 shown in Figure 545.

Figure 547 shows a nucleotide sequence (SEQ ID NO:547) of a native sequence PRO7436 cDNA, wherein SEQ ID NO:547 is a clone designated herein as "DNA119535-2756".

10 Figure 548 shows the amino acid sequence (SEQ ID NO:548) derived from the coding sequence of SEQ ID NO:547 shown in Figure 547.

Figure 549 shows a nucleotide sequence (SEQ ID NO:549) of a native sequence PRO9856 cDNA, wherein SEQ ID NO:549 is a clone designated herein as "DNA119537-2777".

15 Figure 550 shows the amino acid sequence (SEQ ID NO:550) derived from the coding sequence of SEQ ID NO:549 shown in Figure 549.

Figure 551 shows a nucleotide sequence (SEQ ID NO:551) of a native sequence PRO19605 cDNA, wherein SEQ ID NO:551 is a clone designated herein as "DNA119714-2851".

Figure 552 shows the amino acid sequence (SEQ ID NO:552) derived from the coding sequence of SEQ ID NO:551 shown in Figure 551.

20 Figure 553 shows a nucleotide sequence (SEQ ID NO:553) of a native sequence PRO9859 cDNA, wherein SEQ ID NO:553 is a clone designated herein as "DNA125170-2780".

Figure 554 shows the amino acid sequence (SEQ ID NO:554) derived from the coding sequence of SEQ ID NO:553 shown in Figure 553.

25 Figure 555 shows a nucleotide sequence (SEQ ID NO:555) of a native sequence PRO12970 cDNA, wherein SEQ ID NO:555 is a clone designated herein as "DNA129594-2841".

Figure 556 shows the amino acid sequence (SEQ ID NO:556) derived from the coding sequence of SEQ ID NO:555 shown in Figure 555.

Figure 557 shows a nucleotide sequence (SEQ ID NO:557) of a native sequence PRO19626 cDNA, wherein SEQ ID NO:557 is a clone designated herein as "DNA129793-2857".

30 Figure 558 shows the amino acid sequence (SEQ ID NO:558) derived from the coding sequence of SEQ ID NO:557 shown in Figure 557.

Figure 559 shows a nucleotide sequence (SEQ ID NO:559) of a native sequence PRO9833 cDNA, wherein SEQ ID NO:559 is a clone designated herein as "DNA130809-2769".

35 Figure 560 shows the amino acid sequence (SEQ ID NO:560) derived from the coding sequence of SEQ ID NO:559 shown in Figure 559.

Figure 561 shows a nucleotide sequence (SEQ ID NO:561) of a native sequence PRO19670 cDNA, wherein SEQ ID NO:561 is a clone designated herein as "DNA131639-2874".

Figure 562 shows the amino acid sequence (SEQ ID NO:562) derived from the coding sequence of SEQ ID NO:561 shown in Figure 561.

Figure 563 shows a nucleotide sequence (SEQ ID NO:563) of a native sequence PRO19624 cDNA, wherein SEQ ID NO:563 is a clone designated herein as "DNA131649-2855".

5 Figure 564 shows the amino acid sequence (SEQ ID NO:564) derived from the coding sequence of SEQ ID NO:563 shown in Figure 563.

Figure 565 shows a nucleotide sequence (SEQ ID NO:565) of a native sequence PRO19680 cDNA, wherein SEQ ID NO:565 is a clone designated herein as "DNA131652-2876".

Figure 566 shows the amino acid sequence (SEQ ID NO:566) derived from the coding sequence of SEQ ID NO:565 shown in Figure 565.

10 Figure 567 shows a nucleotide sequence (SEQ ID NO:567) of a native sequence PRO19675 cDNA, wherein SEQ ID NO:567 is a clone designated herein as "DNA131658-2875".

Figure 568 shows the amino acid sequence (SEQ ID NO:568) derived from the coding sequence of SEQ ID NO:567 shown in Figure 567.

15 Figure 569 shows a nucleotide sequence (SEQ ID NO:569) of a native sequence PRO9834 cDNA, wherein SEQ ID NO:569 is a clone designated herein as "DNA132162-2770".

Figure 570 shows the amino acid sequence (SEQ ID NO:570) derived from the coding sequence of SEQ ID NO:569 shown in Figure 569.

Figure 571 shows a nucleotide sequence (SEQ ID NO:571) of a native sequence PRO9744 cDNA, wherein SEQ ID NO:571 is a clone designated herein as "DNA136110-2763".

20 Figure 572 shows the amino acid sequence (SEQ ID NO:572) derived from the coding sequence of SEQ ID NO:571 shown in Figure 571.

Figure 573 shows a nucleotide sequence (SEQ ID NO:573) of a native sequence PRO19644 cDNA, wherein SEQ ID NO:573 is a clone designated herein as "DNA139592-2866".

25 Figure 574 shows the amino acid sequence (SEQ ID NO:574) derived from the coding sequence of SEQ ID NO:573 shown in Figure 573.

Figure 575 shows a nucleotide sequence (SEQ ID NO:575) of a native sequence PRO19625 cDNA, wherein SEQ ID NO:575 is a clone designated herein as "DNA139608-2856".

Figure 576 shows the amino acid sequence (SEQ ID NO:576) derived from the coding sequence of SEQ ID NO:575 shown in Figure 575.

30 Figure 577 shows a nucleotide sequence (SEQ ID NO:577) of a native sequence PRO19597 cDNA, wherein SEQ ID NO:577 is a clone designated herein as "DNA143292-2848".

Figure 578 shows the amino acid sequence (SEQ ID NO:578) derived from the coding sequence of SEQ ID NO:577 shown in Figure 577.

35 Figure 579 shows a nucleotide sequence (SEQ ID NO:579) of a native sequence PRO16090 cDNA, wherein SEQ ID NO:579 is a clone designated herein as "DNA144844-2843".

Figure 580 shows the amino acid sequence (SEQ ID NO:580) derived from the coding sequence of SEQ ID NO:579 shown in Figure 579.

Figure 581 shows a nucleotide sequence (SEQ ID NO:581) of a native sequence PRO19576 cDNA, wherein SEQ ID NO:581 is a clone designated herein as "DNA144857-2845".

Figure 582 shows the amino acid sequence (SEQ ID NO:582) derived from the coding sequence of SEQ ID NO:581 shown in Figure 581.

5 Figure 583 shows a nucleotide sequence (SEQ ID NO:583) of a native sequence PRO19646 cDNA, wherein SEQ ID NO:583 is a clone designated herein as "DNA145841-2868".

Figure 584 shows the amino acid sequence (SEQ ID NO:584) derived from the coding sequence of SEQ ID NO:583 shown in Figure 583.

Figure 585 shows a nucleotide sequence (SEQ ID NO:585) of a native sequence PRO19814 cDNA, wherein SEQ ID NO:585 is a clone designated herein as "DNA148004-2882".

10 Figure 586 shows the amino acid sequence (SEQ ID NO:586) derived from the coding sequence of SEQ ID NO:585 shown in Figure 585.

Figure 587 shows a nucleotide sequence (SEQ ID NO:587) of a native sequence PRO19669 cDNA, wherein SEQ ID NO:587 is a clone designated herein as "DNA149893-2873".

15 Figure 588 shows the amino acid sequence (SEQ ID NO:588) derived from the coding sequence of SEQ ID NO:587 shown in Figure 587.

Figure 589 shows a nucleotide sequence (SEQ ID NO:589) of a native sequence PRO19818 cDNA, wherein SEQ ID NO:589 is a clone designated herein as "DNA149930-2884".

Figure 590 shows the amino acid sequence (SEQ ID NO:590) derived from the coding sequence of SEQ ID NO:589 shown in Figure 589.

20 Figure 591 shows a nucleotide sequence (SEQ ID NO:591) of a native sequence PRO20088 cDNA, wherein SEQ ID NO:591 is a clone designated herein as "DNA150157-2898".

Figure 592 shows the amino acid sequence (SEQ ID NO:592) derived from the coding sequence of SEQ ID NO:591 shown in Figure 591.

25 Figure 593 shows a nucleotide sequence (SEQ ID NO:593) of a native sequence PRO16089 cDNA, wherein SEQ ID NO:593 is a clone designated herein as "DNA150163-2842".

Figure 594 shows the amino acid sequence (SEQ ID NO:594) derived from the coding sequence of SEQ ID NO:593 shown in Figure 593.

Figure 595 shows a nucleotide sequence (SEQ ID NO:595) of a native sequence PRO20025 cDNA, wherein SEQ ID NO:595 is a clone designated herein as "DNA153579-2894".

30 Figure 596 shows the amino acid sequence (SEQ ID NO:596) derived from the coding sequence of SEQ ID NO:595 shown in Figure 595.

Figure 597 shows a nucleotide sequence (SEQ ID NO:597) of a native sequence PRO20040 cDNA, wherein SEQ ID NO:597 is a clone designated herein as "DNA164625-2890".

35 Figure 598 shows the amino acid sequence (SEQ ID NO:598) derived from the coding sequence of SEQ ID NO:597 shown in Figure 597.

Figure 599 shows a nucleotide sequence (SEQ ID NO:599) of a native sequence PRO791 cDNA, wherein SEQ ID NO:599 is a clone designated herein as "DNA57838-1337".

Figure 600 shows the amino acid sequence (SEQ ID NO:600) derived from the coding sequence of SEQ ID NO:599 shown in Figure 599.

Figure 601 shows a nucleotide sequence (SEQ ID NO:601) of a native sequence PRO1131 cDNA, wherein SEQ ID NO:601 is a clone designated herein as "DNA59777-1480".

5 Figure 602 shows the amino acid sequence (SEQ ID NO:602) derived from the coding sequence of SEQ ID NO:601 shown in Figure 601.

Figure 603 shows a nucleotide sequence (SEQ ID NO:603) of a native sequence PRO1343 cDNA, wherein SEQ ID NO:603 is a clone designated herein as "DNA66675-1587".

Figure 604 shows the amino acid sequence (SEQ ID NO:604) derived from the coding sequence of SEQ ID NO:603 shown in Figure 603.

10 Figure 605 shows a nucleotide sequence (SEQ ID NO:605) of a native sequence PRO1760 cDNA, wherein SEQ ID NO:605 is a clone designated herein as "DNA76532-1702".

Figure 606 shows the amino acid sequence (SEQ ID NO:606) derived from the coding sequence of SEQ ID NO:605 shown in Figure 605.

15 Figure 607 shows a nucleotide sequence (SEQ ID NO:607) of a native sequence PRO6029 cDNA, wherein SEQ ID NO:607 is a clone designated herein as "DNA105849-2704".

Figure 608 shows the amino acid sequence (SEQ ID NO:608) derived from the coding sequence of SEQ ID NO:607 shown in Figure 607.

Figure 609 shows a nucleotide sequence (SEQ ID NO:609) of a native sequence PRO1801 cDNA, wherein SEQ ID NO:609 is a clone designated herein as "DNA83500-2506".

20 Figure 610 shows the amino acid sequence (SEQ ID NO:610) derived from the coding sequence of SEQ ID NO:609 shown in Figure 609.

## DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

### I. Definitions

25 The terms "PRO polypeptide" and "PRO" as used herein and when immediately followed by a numerical designation refer to various polypeptides, wherein the complete designation (i.e., PRO/number) refers to specific polypeptide sequences as described herein. The terms "PRO/number polypeptide" and "PRO/number" wherein the term "number" is provided as an actual numerical designation as used herein encompass native sequence polypeptides and polypeptide variants (which are further defined herein). The PRO polypeptides described herein  
30 may be isolated from a variety of sources, such as from human tissue types or from another source, or prepared by recombinant or synthetic methods. The term "PRO polypeptide" refers to each individual PRO/number polypeptide disclosed herein. All disclosures in this specification which refer to the "PRO polypeptide" refer to each of the polypeptides individually as well as jointly. For example, descriptions of the preparation of, purification of, derivation of, formation of antibodies to or against, administration of, compositions containing, treatment of a disease with, etc., pertain to each polypeptide of the invention individually. The term "PRO  
35 polypeptide" also includes variants of the PRO/number polypeptides disclosed herein.

A "native sequence PRO polypeptide" comprises a polypeptide having the same amino acid sequence as

the corresponding PRO polypeptide derived from nature. Such native sequence PRO polypeptides can be isolated from nature or can be produced by recombinant or synthetic means. The term "native sequence PRO polypeptide" specifically encompasses naturally-occurring truncated or secreted forms of the specific PRO polypeptide (*e.g.*, an extracellular domain sequence), naturally-occurring variant forms (*e.g.*, alternatively spliced forms) and naturally-occurring allelic variants of the polypeptide. In various embodiments of the invention, the native sequence PRO polypeptides disclosed herein are mature or full-length native sequence polypeptides comprising the full-length amino acids sequences shown in the accompanying figures. Start and stop codons are shown in bold font and underlined in the figures. However, while the PRO polypeptide disclosed in the accompanying figures are shown to begin with methionine residues designated herein as amino acid position 1 in the figures, it is conceivable and possible that other methionine residues located either upstream or downstream from the amino acid position 1 in the figures may be employed as the starting amino acid residue for the PRO polypeptides.

The PRO polypeptide "extracellular domain" or "ECD" refers to a form of the PRO polypeptide which is essentially free of the transmembrane and cytoplasmic domains. Ordinarily, a PRO polypeptide ECD will have less than 1% of such transmembrane and/or cytoplasmic domains and preferably, will have less than 0.5% of such domains. It will be understood that any transmembrane domains identified for the PRO polypeptides of the present invention are identified pursuant to criteria routinely employed in the art for identifying that type of hydrophobic domain. The exact boundaries of a transmembrane domain may vary but most likely by no more than about 5 amino acids at either end of the domain as initially identified herein. Optionally, therefore, an extracellular domain of a PRO polypeptide may contain from about 5 or fewer amino acids on either side of the transmembrane domain/extracellular domain boundary as identified in the Examples or specification and such polypeptides, with or without the associated signal peptide, and nucleic acid encoding them, are contemplated by the present invention.

The approximate location of the "signal peptides" of the various PRO polypeptides disclosed herein are shown in the present specification and/or the accompanying figures. It is noted, however, that the C-terminal boundary of a signal peptide may vary, but most likely by no more than about 5 amino acids on either side of the signal peptide C-terminal boundary as initially identified herein, wherein the C-terminal boundary of the signal peptide may be identified pursuant to criteria routinely employed in the art for identifying that type of amino acid sequence element (*e.g.*, Nielsen et al., Prot. Eng. 10:1-6 (1997) and von Heinje et al., Nucl. Acids. Res. 14:4683-4690 (1986)). Moreover, it is also recognized that, in some cases, cleavage of a signal sequence from a secreted polypeptide is not entirely uniform, resulting in more than one secreted species. These mature polypeptides, where the signal peptide is cleaved within no more than about 5 amino acids on either side of the C-terminal boundary of the signal peptide as identified herein, and the polynucleotides encoding them, are contemplated by the present invention.

"PRO polypeptide variant" means an active PRO polypeptide as defined above or below having at least about 80% amino acid sequence identity with a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Such PRO polypeptide variants include, for instance, PRO

polypeptides wherein one or more amino acid residues are added, or deleted, at the N- or C-terminus of the full-length native amino acid sequence. Ordinarily, a PRO polypeptide variant will have at least about 80% amino acid sequence identity, alternatively at least about 81% amino acid sequence identity, alternatively at least about 82% amino acid sequence identity, alternatively at least about 83% amino acid sequence identity, alternatively at least about 84% amino acid sequence identity, alternatively at least about 85% amino acid sequence identity, alternatively at least about 86% amino acid sequence identity, alternatively at least about 87% amino acid sequence identity, alternatively at least about 88% amino acid sequence identity, alternatively at least about 89% amino acid sequence identity, alternatively at least about 90% amino acid sequence identity, alternatively at least about 91% amino acid sequence identity, alternatively at least about 92% amino acid sequence identity, alternatively at least about 93% amino acid sequence identity, alternatively at least about 94% amino acid sequence identity, alternatively at least about 95% amino acid sequence identity, alternatively at least about 96% amino acid sequence identity, alternatively at least about 97% amino acid sequence identity, alternatively at least about 98% amino acid sequence identity and alternatively at least about 99% amino acid sequence identity to a full-length native sequence PRO polypeptide sequence as disclosed herein, a PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other specifically defined fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, PRO variant polypeptides are at least about 10 amino acids in length, alternatively at least about 20 amino acids in length, alternatively at least about 30 amino acids in length, alternatively at least about 40 amino acids in length, alternatively at least about 50 amino acids in length, alternatively at least about 60 amino acids in length, alternatively at least about 70 amino acids in length, alternatively at least about 80 amino acids in length, alternatively at least about 90 amino acids in length, alternatively at least about 100 amino acids in length, alternatively at least about 150 amino acids in length, alternatively at least about 200 amino acids in length, alternatively at least about 300 amino acids in length, or more.

"Percent (%) amino acid sequence identity" with respect to the PRO polypeptide sequences identified herein is defined as the percentage of amino acid residues in a candidate sequence that are identical with the amino acid residues in the specific PRO polypeptide sequence, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity, and not considering any conservative substitutions as part of the sequence identity. Alignment for purposes of determining percent amino acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. Those skilled in the art can determine appropriate parameters for measuring alignment, including any algorithms needed to achieve maximal alignment over the full length of the sequences being compared. For purposes herein, however, % amino acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through

Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A. As examples of % amino acid sequence identity calculations using this method, Tables 2 and 3 demonstrate how to calculate the % amino acid sequence identity of the amino acid sequence designated "Comparison Protein" to the amino acid sequence designated "PRO", wherein "PRO" represents the amino acid sequence of a hypothetical PRO polypeptide of interest, "Comparison Protein" represents the amino acid sequence of a polypeptide against which the "PRO" polypeptide of interest is being compared, and "X", "Y" and "Z" each represent different hypothetical amino acid residues.

Unless specifically stated otherwise, all % amino acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % amino acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % amino acid sequence identity value is determined by dividing (a) the number of matching identical amino acid residues between the amino acid sequence of the PRO polypeptide of interest having a sequence derived from the native PRO polypeptide and the comparison amino acid sequence of interest (i.e., the sequence against which the PRO polypeptide of interest is being compared which may be a PRO variant polypeptide) as determined by WU-BLAST-2 by (b) the total number of amino acid residues of the PRO polypeptide of interest. For example, in the statement "a polypeptide comprising an the amino acid sequence A which has or having at least 80% amino acid sequence identity to the amino acid sequence B", the amino acid sequence A is the comparison amino acid sequence of interest and the amino acid sequence B is the amino acid sequence of the PRO polypeptide of interest.

Percent amino acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the

National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for amino acid sequence comparisons, the % amino acid sequence identity of a given amino acid sequence A to, with, or against a given amino acid sequence B (which can alternatively be phrased as a given amino acid sequence A that has or comprises a certain % amino acid sequence identity to, with, or against a given amino acid sequence B) is calculated as follows:

$$100 \text{ times the fraction } X/Y$$

where X is the number of amino acid residues scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of A and B, and where Y is the total number of amino acid residues in B. It will be appreciated that where the length of amino acid sequence A is not equal to the length of amino acid sequence B, the % amino acid sequence identity of A to B will not equal the % amino acid sequence identity of B to A.

"PRO variant polynucleotide" or "PRO variant nucleic acid sequence" means a nucleic acid molecule which encodes an active PRO polypeptide as defined below and which has at least about 80% nucleic acid sequence identity with a nucleotide acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal peptide, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Ordinarily, a PRO variant polynucleotide will have at least about 80% nucleic acid sequence identity, alternatively at least about 81% nucleic acid sequence identity, alternatively at least about 82% nucleic acid sequence identity, alternatively at least about 83% nucleic acid sequence identity, alternatively at least about 84% nucleic acid sequence identity, alternatively at least about 85% nucleic acid sequence identity, alternatively at least about 86% nucleic acid sequence identity, alternatively at least about 87% nucleic acid sequence identity, alternatively at least about 88% nucleic acid sequence identity, alternatively at least about 89% nucleic acid sequence identity, alternatively at least about 90% nucleic acid sequence identity, alternatively at least about 91% nucleic acid sequence identity, alternatively at least about 92% nucleic acid sequence identity, alternatively at least about 93% nucleic acid sequence identity, alternatively at least about 94% nucleic acid sequence identity, alternatively at least about 95% nucleic acid sequence identity, alternatively at least about 96% nucleic acid sequence identity, alternatively at least about 97% nucleic acid sequence identity, alternatively at least about 98% nucleic acid sequence identity and alternatively at least about 99% nucleic acid sequence identity with a nucleic acid sequence encoding a full-length native sequence PRO polypeptide sequence as disclosed herein, a full-length native sequence PRO polypeptide sequence lacking the signal peptide as disclosed herein, an extracellular domain of a PRO polypeptide, with or without the signal sequence, as disclosed herein or any other fragment of a full-length PRO polypeptide sequence as disclosed herein. Variants do not encompass the native nucleotide sequence.



Ordinarily, PRO variant polynucleotides are at least about 30 nucleotides in length, alternatively at least about 60 nucleotides in length, alternatively at least about 90 nucleotides in length, alternatively at least about 120 nucleotides in length, alternatively at least about 150 nucleotides in length, alternatively at least about 180 nucleotides in length, alternatively at least about 210 nucleotides in length, alternatively at least about 240 nucleotides in length, alternatively at least about 270 nucleotides in length, alternatively at least about 300 nucleotides in length, alternatively at least about 450 nucleotides in length, alternatively at least about 600 nucleotides in length, alternatively at least about 900 nucleotides in length, or more.

"Percent (%) nucleic acid sequence identity" with respect to PRO-encoding nucleic acid sequences identified herein is defined as the percentage of nucleotides in a candidate sequence that are identical with the nucleotides in the PRO nucleic acid sequence of interest, after aligning the sequences and introducing gaps, if necessary, to achieve the maximum percent sequence identity. Alignment for purposes of determining percent nucleic acid sequence identity can be achieved in various ways that are within the skill in the art, for instance, using publicly available computer software such as BLAST, BLAST-2, ALIGN or Megalign (DNASTAR) software. For purposes herein, however, % nucleic acid sequence identity values are generated using the sequence comparison computer program ALIGN-2, wherein the complete source code for the ALIGN-2 program is provided in Table 1 below. The ALIGN-2 sequence comparison computer program was authored by Genentech, Inc. and the source code shown in Table 1 below has been filed with user documentation in the U.S. Copyright Office, Washington D.C., 20559, where it is registered under U.S. Copyright Registration No. TXU510087. The ALIGN-2 program is publicly available through Genentech, Inc., South San Francisco, California or may be compiled from the source code provided in Table 1 below. The ALIGN-2 program should be compiled for use on a UNIX operating system, preferably digital UNIX V4.0D. All sequence comparison parameters are set by the ALIGN-2 program and do not vary.

In situations where ALIGN-2 is employed for nucleic acid sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

$$100 \text{ times the fraction } W/Z$$

where W is the number of nucleotides scored as identical matches by the sequence alignment program ALIGN-2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C. As examples of % nucleic acid sequence identity calculations, Tables 4 and 5, demonstrate how to calculate the % nucleic acid sequence identity of the nucleic acid sequence designated "Comparison DNA" to the nucleic acid sequence designated "PRO-DNA", wherein "PRO-DNA" represents a hypothetical PRO-encoding nucleic acid sequence of interest, "Comparison DNA" represents the nucleotide sequence of a nucleic acid molecule against which the "PRO-DNA" nucleic acid molecule of interest is being compared, and "N", "L" and "V" each represent

different hypothetical nucleotides.

Unless specifically stated otherwise, all % nucleic acid sequence identity values used herein are obtained as described in the immediately preceding paragraph using the ALIGN-2 computer program. However, % nucleic acid sequence identity values may also be obtained as described below by using the WU-BLAST-2 computer program (Altschul et al., Methods in Enzymology 266:460-480 (1996)). Most of the WU-BLAST-2 search parameters are set to the default values. Those not set to default values, i.e., the adjustable parameters, are set with the following values: overlap span = 1, overlap fraction = 0.125, word threshold (T) = 11, and scoring matrix = BLOSUM62. When WU-BLAST-2 is employed, a % nucleic acid sequence identity value is determined by dividing (a) the number of matching identical nucleotides between the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest having a sequence derived from the native sequence PRO polypeptide-encoding nucleic acid and the comparison nucleic acid molecule of interest (i.e., the sequence against which the PRO polypeptide-encoding nucleic acid molecule of interest is being compared which may be a variant PRO polynucleotide) as determined by WU-BLAST-2 by (b) the total number of nucleotides of the PRO polypeptide-encoding nucleic acid molecule of interest. For example, in the statement "an isolated nucleic acid molecule comprising a nucleic acid sequence A which has or having at least 80% nucleic acid sequence identity to the nucleic acid sequence B", the nucleic acid sequence A is the comparison nucleic acid molecule of interest and the nucleic acid sequence B is the nucleic acid sequence of the PRO polypeptide-encoding nucleic acid molecule of interest.

Percent nucleic acid sequence identity may also be determined using the sequence comparison program NCBI-BLAST2 (Altschul et al., Nucleic Acids Res. 25:3389-3402 (1997)). The NCBI-BLAST2 sequence comparison program may be downloaded from <http://www.ncbi.nlm.nih.gov> or otherwise obtained from the National Institute of Health, Bethesda, MD. NCBI-BLAST2 uses several search parameters, wherein all of those search parameters are set to default values including, for example, unmask = yes, strand = all, expected occurrences = 10, minimum low complexity length = 15/5, multi-pass e-value = 0.01, constant for multi-pass = 25, dropoff for final gapped alignment = 25 and scoring matrix = BLOSUM62.

In situations where NCBI-BLAST2 is employed for sequence comparisons, the % nucleic acid sequence identity of a given nucleic acid sequence C to, with, or against a given nucleic acid sequence D (which can alternatively be phrased as a given nucleic acid sequence C that has or comprises a certain % nucleic acid sequence identity to, with, or against a given nucleic acid sequence D) is calculated as follows:

100 times the fraction  $W/Z$

where W is the number of nucleotides scored as identical matches by the sequence alignment program NCBI-BLAST2 in that program's alignment of C and D, and where Z is the total number of nucleotides in D. It will be appreciated that where the length of nucleic acid sequence C is not equal to the length of nucleic acid sequence D, the % nucleic acid sequence identity of C to D will not equal the % nucleic acid sequence identity of D to C.

In other embodiments, PRO variant polynucleotides are nucleic acid molecules that encode an active PRO polypeptide and which are capable of hybridizing, preferably under stringent hybridization and wash conditions,

to nucleotide sequences encoding a full-length PRO polypeptide as disclosed herein. PRO variant polypeptides may be those that are encoded by a PRO variant polynucleotide.

5 "Isolated," when used to describe the various polypeptides disclosed herein, means polypeptide that has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials that would typically interfere with diagnostic or therapeutic  
10 uses for the polypeptide, and may include enzymes, hormones, and other proteinaceous or non-proteinaceous solutes. In preferred embodiments, the polypeptide will be purified (1) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence by use of a spinning cup sequenator, or (2) to homogeneity by SDS-PAGE under non-reducing or reducing conditions using Coomassie blue or, preferably, silver stain. Isolated polypeptide includes polypeptide *in situ* within recombinant cells, since at least one  
15 component of the PRO polypeptide natural environment will not be present. Ordinarily, however, isolated polypeptide will be prepared by at least one purification step.

An "isolated" PRO polypeptide-encoding nucleic acid or other polypeptide-encoding nucleic acid is a nucleic acid molecule that is identified and separated from at least one contaminant nucleic acid molecule with which it is ordinarily associated in the natural source of the polypeptide-encoding nucleic acid. An isolated  
20 polypeptide-encoding nucleic acid molecule is other than in the form or setting in which it is found in nature. Isolated polypeptide-encoding nucleic acid molecules therefore are distinguished from the specific polypeptide-encoding nucleic acid molecule as it exists in natural cells. However, an isolated polypeptide-encoding nucleic acid molecule includes polypeptide-encoding nucleic acid molecules contained in cells that ordinarily express the polypeptide where, for example, the nucleic acid molecule is in a chromosomal location different from that of  
25 natural cells.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

25 Nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; or a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked"  
30 means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

The term "antibody" is used in the broadest sense and specifically covers, for example, single anti-PRO  
35 monoclonal antibodies (including agonist, antagonist, and neutralizing antibodies), anti-PRO antibody compositions with polypeptopic specificity, single chain anti-PRO antibodies, and fragments of anti-PRO antibodies (see below). The term "monoclonal antibody" as used herein refers to an antibody obtained from a population of substantially

homogeneous antibodies, i.e., the individual antibodies comprising the population are identical except for possible naturally-occurring mutations that may be present in minor amounts.

"Stringency" of hybridization reactions is readily determinable by one of ordinary skill in the art, and generally is an empirical calculation dependent upon probe length, washing temperature, and salt concentration. In general, longer probes require higher temperatures for proper annealing, while shorter probes need lower temperatures. Hybridization generally depends on the ability of denatured DNA to reanneal when complementary strands are present in an environment below their melting temperature. The higher the degree of desired homology between the probe and hybridizable sequence, the higher the relative temperature which can be used. As a result, it follows that higher relative temperatures would tend to make the reaction conditions more stringent, while lower temperatures less so. For additional details and explanation of stringency of hybridization reactions, see Ausubel et al., Current Protocols in Molecular Biology, Wiley Interscience Publishers, (1995).

"Stringent conditions" or "high stringency conditions", as defined herein, may be identified by those that: (1) employ low ionic strength and high temperature for washing, for example 0.015 M sodium chloride/0.0015 M sodium citrate/0.1% sodium dodecyl sulfate at 50°C; (2) employ during hybridization a denaturing agent, such as formamide, for example, 50% (v/v) formamide with 0.1% bovine serum albumin/0.1% Ficoll/0.1% polyvinylpyrrolidone/50mM sodium phosphate buffer at pH 6.5 with 750 mM sodium chloride, 75 mM sodium citrate at 42°C; or (3) employ 50% formamide, 5 x SSC (0.75 M NaCl, 0.075 M sodium citrate), 50 mM sodium phosphate (pH 6.8), 0.1% sodium pyrophosphate, 5 x Denhardt's solution, sonicated salmon sperm DNA (50 µg/ml), 0.1% SDS, and 10% dextran sulfate at 42°C, with washes at 42°C in 0.2 x SSC (sodium chloride/sodium citrate) and 50% formamide at 55°C, followed by a high-stringency wash consisting of 0.1 x SSC containing EDTA at 55°C.

"Moderately stringent conditions" may be identified as described by Sambrook et al., Molecular Cloning: A Laboratory Manual, New York: Cold Spring Harbor Press, 1989, and include the use of washing solution and hybridization conditions (e.g., temperature, ionic strength and %SDS) less stringent than those described above. An example of moderately stringent conditions is overnight incubation at 37°C in a solution comprising: 20% formamide, 5 x SSC (150 mM NaCl, 15 mM trisodium citrate), 50 mM sodium phosphate (pH 7.6), 5 x Denhardt's solution, 10% dextran sulfate, and 20 mg/ml denatured sheared salmon sperm DNA, followed by washing the filters in 1 x SSC at about 37-50°C. The skilled artisan will recognize how to adjust the temperature, ionic strength, etc. as necessary to accommodate factors such as probe length and the like.

The term "epitope tagged" when used herein refers to a chimeric polypeptide comprising a PRO polypeptide fused to a "tag polypeptide". The tag polypeptide has enough residues to provide an epitope against which an antibody can be made, yet is short enough such that it does not interfere with activity of the polypeptide to which it is fused. The tag polypeptide preferably also is fairly unique so that the antibody does not substantially cross-react with other epitopes. Suitable tag polypeptides generally have at least six amino acid residues and usually between about 8 and 50 amino acid residues (preferably, between about 10 and 20 amino acid residues).

As used herein, the term "immunoadhesin" designates antibody-like molecules which combine the binding specificity of a heterologous protein (an "adhesin") with the effector functions of immunoglobulin constant domains. Structurally, the immunoadhesins comprise a fusion of an amino acid sequence with the desired binding

specificity which is other than the antigen recognition and binding site of an antibody (i.e., is "heterologous"), and an immunoglobulin constant domain sequence. The adhesin part of an immunoadhesin molecule typically is a contiguous amino acid sequence comprising at least the binding site of a receptor or a ligand. The immunoglobulin constant domain sequence in the immunoadhesin may be obtained from any immunoglobulin, such as IgG-1, IgG-2, IgG-3, or IgG-4 subtypes, IgA (including IgA-1 and IgA-2), IgE, IgD or IgM.

5 "Active" or "activity" for the purposes herein refers to form(s) of a PRO polypeptide which retain a biological and/or an immunological activity of native or naturally-occurring PRO, wherein "biological" activity refers to a biological function (either inhibitory or stimulatory) caused by a native or naturally-occurring PRO other than the ability to induce the production of an antibody against an antigenic epitope possessed by a native or naturally-occurring PRO and an "immunological" activity refers to the ability to induce the production of an  
10 antibody against an antigenic epitope possessed by a native or naturally-occurring PRO.

The term "antagonist" is used in the broadest sense, and includes any molecule that partially or fully blocks, inhibits, or neutralizes a biological activity of a native PRO polypeptide disclosed herein. In a similar manner, the term "agonist" is used in the broadest sense and includes any molecule that mimics a biological activity of a native PRO polypeptide disclosed herein. Suitable agonist or antagonist molecules specifically  
15 include agonist or antagonist antibodies or antibody fragments, fragments or amino acid sequence variants of native PRO polypeptides, peptides, antisense oligonucleotides, small organic molecules, etc. Methods for identifying agonists or antagonists of a PRO polypeptide may comprise contacting a PRO polypeptide with a candidate agonist or antagonist molecule and measuring a detectable change in one or more biological activities normally associated with the PRO polypeptide.

20 "Treatment" refers to both therapeutic treatment and prophylactic or preventative measures, wherein the object is to prevent or slow down (lessen) the targeted pathologic condition or disorder. Those in need of treatment include those already with the disorder as well as those prone to have the disorder or those in whom the disorder is to be prevented.

"Chronic" administration refers to administration of the agent(s) in a continuous mode as opposed to an  
25 acute mode, so as to maintain the initial therapeutic effect (activity) for an extended period of time. "Intermittent" administration is treatment that is not consecutively done without interruption, but rather is cyclic in nature.

"Mammal" for purposes of treatment refers to any animal classified as a mammal, including humans, domestic and farm animals, and zoo, sports, or pet animals, such as dogs, cats, cattle, horses, sheep, pigs, goats, rabbits, etc. Preferably, the mammal is human.

30 Administration "in combination with" one or more further therapeutic agents includes simultaneous (concurrent) and consecutive administration in any order.

"Carriers" as used herein include pharmaceutically acceptable carriers, excipients, or stabilizers which are nontoxic to the cell or mammal being exposed thereto at the dosages and concentrations employed. Often the physiologically acceptable carrier is an aqueous pH buffered solution. Examples of physiologically acceptable  
35 carriers include buffers such as phosphate, citrate, and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptide; proteins, such as serum albumin, gelatin, or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone; amino acids such as glycine, glutamine,

asparagine, arginine or lysine; monosaccharides, disaccharides, and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN™, polyethylene glycol (PEG), and PLURONICS™.

5 "Antibody fragments" comprise a portion of an intact antibody, preferably the antigen binding or variable region of the intact antibody. Examples of antibody fragments include Fab, Fab', F(ab')<sub>2</sub>, and Fv fragments; diabodies; linear antibodies (Zapata et al., Protein Eng. 8(10): 1057-1062 [1995]); single-chain antibody molecules; and multispecific antibodies formed from antibody fragments.

10 Papain digestion of antibodies produces two identical antigen-binding fragments, called "Fab" fragments, each with a single antigen-binding site, and a residual "Fc" fragment, a designation reflecting the ability to crystallize readily. Pepsin treatment yields an F(ab')<sub>2</sub> fragment that has two antigen-combining sites and is still capable of cross-linking antigen.

"Fv" is the minimum antibody fragment which contains a complete antigen-recognition and -binding site. This region consists of a dimer of one heavy- and one light-chain variable domain in tight, non-covalent association. It is in this configuration that the three CDRs of each variable domain interact to define an antigen-binding site on the surface of the V<sub>H</sub>-V<sub>L</sub> dimer. Collectively, the six CDRs confer antigen-binding specificity to the antibody. However, even a single variable domain (or half of an Fv comprising only three CDRs specific for an antigen) has the ability to recognize and bind antigen, although at a lower affinity than the entire binding site.

15 The Fab fragment also contains the constant domain of the light chain and the first constant domain (CH1) of the heavy chain. Fab fragments differ from Fab' fragments by the addition of a few residues at the carboxy terminus of the heavy chain CH1 domain including one or more cysteines from the antibody hinge region. Fab'-SH is the designation herein for Fab' in which the cysteine residue(s) of the constant domains bear a free thiol group. F(ab')<sub>2</sub> antibody fragments originally were produced as pairs of Fab' fragments which have hinge cysteines between them. Other chemical couplings of antibody fragments are also known.

20 The "light chains" of antibodies (immunoglobulins) from any vertebrate species can be assigned to one of two clearly distinct types, called kappa and lambda, based on the amino acid sequences of their constant domains.

Depending on the amino acid sequence of the constant domain of their heavy chains, immunoglobulins can be assigned to different classes. There are five major classes of immunoglobulins: IgA, IgD, IgE, IgG, and IgM, and several of these may be further divided into subclasses (isotypes), e.g., IgG1, IgG2, IgG3, IgG4, IgA, and IgA2.

30 "Single-chain Fv" or "sFv" antibody fragments comprise the V<sub>H</sub> and V<sub>L</sub> domains of antibody, wherein these domains are present in a single polypeptide chain. Preferably, the Fv polypeptide further comprises a polypeptide linker between the V<sub>H</sub> and V<sub>L</sub> domains which enables the sFv to form the desired structure for antigen binding. For a review of sFv, see Pluckthun in The Pharmacology of Monoclonal Antibodies, vol. 113, Rosenberg and Moore eds., Springer-Verlag, New York, pp. 269-315 (1994).

35 The term "diabodies" refers to small antibody fragments with two antigen-binding sites, which fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) in the same

polypeptide chain ( $V_H$ - $V_L$ ). By using a linker that is too short to allow pairing between the two domains on the same chain, the domains are forced to pair with the complementary domains of another chain and create two antigen-binding sites. Diabodies are described more fully in, for example, EP 404,097; WO 93/11161; and Hollinger et al., Proc. Natl. Acad. Sci. USA, 90:6444-6448 (1993).

5 An "isolated" antibody is one which has been identified and separated and/or recovered from a component of its natural environment. Contaminant components of its natural environment are materials which would interfere with diagnostic or therapeutic uses for the antibody, and may include enzymes, hormones, and other proteinaceous or nonproteinaceous solutes. In preferred embodiments, the antibody will be purified (1) to greater than 95% by weight of antibody as determined by the Lowry method, and most preferably more than 99% by weight, (2) to a degree sufficient to obtain at least 15 residues of N-terminal or internal amino acid sequence  
10 by use of a spinning cup sequenator, or (3) to homogeneity by SDS-PAGE under reducing or nonreducing conditions using Coomassie blue or, preferably, silver stain. Isolated antibody includes the antibody in situ within recombinant cells since at least one component of the antibody's natural environment will not be present. Ordinarily, however, isolated antibody will be prepared by at least one purification step.

15 An antibody that "specifically binds to" or is "specific for" a particular polypeptide or an epitope on a particular polypeptide is one that binds to that particular polypeptide or epitope on a particular polypeptide without substantially binding to any other polypeptide or polypeptide epitope.

The word "label" when used herein refers to a detectable compound or composition which is conjugated directly or indirectly to the antibody so as to generate a "labeled" antibody. The label may be detectable by itself (e.g. radioisotope labels or fluorescent labels) or, in the case of an enzymatic label, may catalyze chemical  
20 alteration of a substrate compound or composition which is detectable.

By "solid phase" is meant a non-aqueous matrix to which the antibody of the present invention can adhere. Examples of solid phases encompassed herein include those formed partially or entirely of glass (e.g., controlled pore glass), polysaccharides (e.g., agarose), polyacrylamides, polystyrene, polyvinyl alcohol and silicones. In certain embodiments, depending on the context, the solid phase can comprise the well of an assay  
25 plate; in others it is a purification column (e.g., an affinity chromatography column). This term also includes a discontinuous solid phase of discrete particles, such as those described in U.S. Patent No. 4,275,149.

A "liposome" is a small vesicle composed of various types of lipids, phospholipids and/or surfactant which is useful for delivery of a drug (such as a PRO polypeptide or antibody thereto) to a mammal. The components of the liposome are commonly arranged in a bilayer formation, similar to the lipid arrangement of  
30 biological membranes.

A "small molecule" is defined herein to have a molecular weight below about 500 Daltons.

An "effective amount" of a polypeptide disclosed herein or an agonist or antagonist thereof is an amount sufficient to carry out a specifically stated purpose. An "effective amount" may be determined empirically and in a routine manner, in relation to the stated purpose.  
35

**Table 1**

```

/*
*
* C-C increased from 12 to 15
* Z is average of EQ
5  * B is average of ND
* match with stop is _M; stop-stop = 0; J (joker) match = 0
*/
#define _M      -8      /* value of a match with a stop */

10 int _day[26][26] = {
/*  A B C D E F G H I J K L M N O P Q R S T U V W X Y Z */
/* A */ { 2, 0, -2, 0, 0, -4, 1, -1, -1, 0, -1, -2, -1, 0, _M, 1, 0, -2, 1, 1, 0, 0, -6, 0, -3, 0},
/* B */ { 0, 3, -4, 3, 2, -5, 0, 1, -2, 0, 0, -3, -2, 2, _M, -1, 1, 0, 0, 0, 0, -2, -5, 0, -3, 1},
/* C */ {-2, -4, 15, -5, -5, -4, -3, -3, -2, 0, -5, -6, -5, -4, _M, -3, -5, -4, 0, -2, 0, -2, -8, 0, 0, -5},
15 /* D */ { 0, 3, -5, 4, 3, -6, 1, 1, -2, 0, 0, -4, -3, 2, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 2},
/* E */ { 0, 2, -5, 3, 4, -5, 0, 1, -2, 0, 0, -3, -2, 1, _M, -1, 2, -1, 0, 0, 0, -2, -7, 0, -4, 3},
/* F */ {-4, -5, -4, -6, -5, 9, -5, -2, 1, 0, -5, 2, 0, -4, _M, -5, -5, -4, -3, -3, 0, -1, 0, 0, 7, -5},
/* G */ { 1, 0, -3, 1, 0, -5, 5, -2, -3, 0, -2, -4, -3, 0, _M, -1, -1, -3, 1, 0, 0, -1, -7, 0, -5, 0},
/* H */ {-1, 1, -3, 1, 1, -2, -2, 6, -2, 0, 0, -2, -2, 2, _M, 0, 3, 2, -1, -1, 0, -2, -3, 0, 0, 2},
20 /* I */ {-1, -2, -2, -2, -2, 1, -3, -2, 5, 0, -2, 2, 2, -2, _M, -2, -2, -2, -1, 0, 0, 4, -5, 0, -1, -2},
/* J */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* K */ {-1, 0, -5, 0, 0, -5, -2, 0, -2, 0, 5, -3, 0, 1, _M, -1, 1, 3, 0, 0, 0, -2, -3, 0, -4, 0},
/* L */ {-2, -3, -6, -4, -3, 2, -4, -2, 2, 0, -3, 6, 4, -3, _M, -3, -2, -3, -3, -1, 0, 2, -2, 0, -1, -2},
/* M */ {-1, -2, -5, -3, -2, 0, -3, -2, 2, 0, 0, 4, 6, -2, _M, -2, -1, 0, -2, -1, 0, 2, -4, 0, -2, -1},
25 /* N */ { 0, 2, -4, 2, 1, -4, 0, 2, -2, 0, 1, -3, -2, 2, _M, -1, 1, 0, 1, 0, 0, -2, -4, 0, -2, 1},
/* O */ { _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, _M, 0, _M, _M, _M, _M, _M, _M, _M, _M, _M},
/* P */ { 1, -1, -3, -1, -1, -5, -1, 0, -2, 0, -1, -3, -2, -1, _M, 6, 0, 0, 1, 0, 0, -1, -6, 0, -5, 0},
/* Q */ { 0, 1, -5, 2, 2, -5, -1, 3, -2, 0, 1, -2, -1, 1, _M, 0, 4, 1, -1, -1, 0, -2, -5, 0, -4, 3},
/* R */ {-2, 0, -4, -1, -1, -4, -3, 2, -2, 0, 3, -3, 0, 0, _M, 0, 1, 6, 0, -1, 0, -2, 2, 0, -4, 0},
30 /* S */ { 1, 0, 0, 0, 0, -3, 1, -1, -1, 0, 0, -3, -2, 1, _M, 1, -1, 0, 2, 1, 0, -1, -2, 0, -3, 0},
/* T */ { 1, 0, -2, 0, 0, -3, 0, -1, 0, 0, 0, -1, -1, 0, _M, 0, -1, -1, 1, 3, 0, 0, -5, 0, -3, 0},
/* U */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* V */ { 0, -2, -2, -2, -2, -1, -1, -2, 4, 0, -2, 2, 2, -2, _M, -1, -2, -2, -1, 0, 0, 4, -6, 0, -2, -2},
/* W */ {-6, -5, -8, -7, -7, 0, -7, -3, -5, 0, -3, -2, -4, -4, _M, -6, -5, 2, -2, -5, 0, -6, 17, 0, 0, -6},
35 /* X */ { 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, 0, _M, 0, 0, 0, 0, 0, 0, 0, 0, 0},
/* Y */ {-3, -3, 0, -4, -4, 7, -5, 0, -1, 0, -4, -1, -2, -2, _M, -5, -4, -4, -3, -3, 0, -2, 0, 0, 10, -4},
/* Z */ { 0, 1, -5, 2, 3, -5, 0, 2, -2, 0, 0, -2, -1, 1, _M, 0, 3, 0, 0, 0, 0, -2, -6, 0, -4, 4}
};

40

45

50

55

```



**Table 1 (cont')**

```

/*
*/
#include <stdio.h>
#include <ctype.h>

5
#define MAXJMP      16      /* max jumps in a diag */
#define MAXGAP      24      /* don't continue to penalize gaps larger than this */
#define JMPS        1024    /* max jmps in an path */
10
#define MX          4       /* save if there's at least MX-1 bases since last jmp */

#define DMAT        3       /* value of matching bases */
#define DMIS        0       /* penalty for mismatched bases */
#define DINS0       8       /* penalty for a gap */
15
#define DINS1       1       /* penalty per base */
#define PINS0       8       /* penalty for a gap */
#define PINS1       4       /* penalty per residue */

struct jmp {
20
    short            n[MAXJMP]; /* size of jmp (neg for dely) */
    unsigned short   x[MAXJMP]; /* base no. of jmp in seq x */
}; /* limits seq to 2^16 -1 */

struct diag {
25
    int              score;      /* score at last jmp */
    long             offset;     /* offset of prev block */
    short            jmp;        /* current jmp index */
    struct jmp        jp;        /* list of jmps */
};

30
struct path {
    int              spc;         /* number of leading spaces */
    short            n[JMPS]; /* size of jmp (gap) */
    int              x[JMPS]; /* loc of jmp (last elem before gap) */
35
};

char            *ofile;          /* output file name */
char            *namex[2];       /* seq names: getseqs() */
char            *prog;           /* prog name for err msgs */
40
char            *seqx[2];        /* seqs: getseqs() */
int             dmax;            /* best diag: nw() */
int             dmax0;          /* final diag */
int             dna;            /* set if dna: main() */
int             endgaps;        /* set if penalizing end gaps */
45
int             gapx, gapy;      /* total gaps in seqs */
int             len0, len1;      /* seq lens */
int             ngapx, ngapy;    /* total size of gaps */
int             smax;           /* max score: nw() */
int             *xbm;           /* bitmap for matching */
50
long            offset;         /* current offset in jmp file */
struct          diag            /* holds diagonals */
struct          path            /* holds path for seqs */
    pp[2];

char            *calloc(), *malloc(), *index(), *strcpy();
55
char            *getseq(), *g_calloc();

```

60

**Table 1 (cont')**

```

/* Needleman-Wunsch alignment program
*
* usage: progs file1 file2
* where file1 and file2 are two dna or two protein sequences.
5 * The sequences can be in upper- or lower-case and may contain ambiguity
* Any lines beginning with ';', '>' or '<' are ignored
* Max file length is 65535 (limited by unsigned short x in the jmp struct)
* A sequence with 1/3 or more of its elements ACGTU is assumed to be DNA
10 * Output is in the file "align.out"
*
* The program may create a tmp file in /tmp to hold info about traceback.
* Original version developed under BSD 4.3 on a vax 8650
*/
15 #include "nw.h"
#include "day.h"

static _dbval[26] = {
    1,14,2,13,0,0,4,11,0,0,12,0,3,15,0,0,0,5,6,8,8,7,9,0,10,0
20 };

static _pbval[26] = {
    1, 2|(1<<('D'-'A'))|(1<<('N'-'A')), 4, 8, 16, 32, 64,
    128, 256, 0xFFFFFFFF, 1<<10, 1<<11, 1<<12, 1<<13, 1<<14,
    1<<15, 1<<16, 1<<17, 1<<18, 1<<19, 1<<20, 1<<21, 1<<22,
25 1<<23, 1<<24, 1<<25|(1<<('E'-'A'))|(1<<('Q'-'A'))
};

main(ac, av)
30     int      ac;
    char      *av[];
{
    prog = av[0];
    if (ac != 3) {
35         fprintf(stderr, "usage: %s file1 file2\n", prog);
        fprintf(stderr, "where file1 and file2 are two dna or two protein sequences.\n");
        fprintf(stderr, "The sequences can be in upper- or lower-case\n");
        fprintf(stderr, "Any lines beginning with ';' or '<' are ignored\n");
        fprintf(stderr, "Output is in the file \"align.out\"\n");
        exit(1);
40     }
    namex[0] = av[1];
    namex[1] = av[2];
    seqx[0] = getseq(namex[0], &len0);
    seqx[1] = getseq(namex[1], &len1);
45     xbm = (dna)? _dbval : _pbval;

    endgaps = 0; /* 1 to penalize endgaps */
    ofile = "align.out"; /* output file */

50     nw(); /* fill in the matrix, get the possible jumps */
    readjumps(); /* get the actual jumps */
    print(); /* print stats, alignment */

    cleanup(0); /* unlink any tmp files */
55 }

```

**main**

**Table 1 (cont')**

```

/* do the alignment, return best score: main()
* dna: values in Fitch and Smith, PNAS, 80, 1382-1386, 1983
* pro: PAM 250 values
5  * When scores are equal, we prefer mismatches to any gap, prefer
* a new gap to extending an ongoing gap, and prefer a gap in seqx
* to a gap in seq y.
*/
nw()
{
10  char      *px, *py;          /* seqs and ptrs */
    int      *ndely, *dely;     /* keep track of dely */
    int      ndelx, delx;       /* keep track of delx */
    int      *tmp;              /* for swapping row0, row1 */
    int      mis;               /* score for each type */
15  int      ins0, ins1;         /* insertion penalties */
    register id;                 /* diagonal index */
    register ij;                 /* jmp index */
    register *col0, *col1;       /* score for curr, last row */
    register xx, yy;             /* index into seqs */
20
    dx = (struct diag *)g_calloc("to get diags", len0+len1+1, sizeof(struct diag));

    ndely = (int *)g_calloc("to get ndely", len1+1, sizeof(int));
    dely = (int *)g_calloc("to get dely", len1+1, sizeof(int));
25  col0 = (int *)g_calloc("to get col0", len1+1, sizeof(int));
    col1 = (int *)g_calloc("to get col1", len1+1, sizeof(int));
    ins0 = (dna)? DINS0 : PINS0;
    ins1 = (dna)? DINS1 : PINS1;

30  smax = -10000;
    if (endgaps) {
        for (col0[0] = dely[0] = -ins0, yy = 1; yy <= len1; yy++) {
            col0[yy] = dely[yy] = col0[yy-1] - ins1;
            ndely[yy] = yy;
35  }
        col0[0] = 0;          /* Waterman Bull Math Biol 84 */
    }
    else
40  for (yy = 1; yy <= len1; yy++)
        dely[yy] = -ins0;

    /* fill in match matrix
    */
45  for (px = seqx[0], xx = 1; xx <= len0; px++, xx++) {
        /* initialize first entry in col
        */
        if (endgaps) {
            if (xx == 1)
                col1[0] = delx = -(ins0+ins1);
50  else
                col1[0] = delx = col0[0] - ins1;
            ndelx = xx;
        }
        else {
55  col1[0] = 0;
            delx = -ins0;
            ndelx = 0;
        }
60

```

**Table 1 (cont')****...nw**

```

5      for (py = seqx[1], yy = 1; yy <= len1; py++, yy++) {
        mis = col0[yy-1];
        if (dna)
            mis += (xbm[*px-'A']&xbm[*py-'A'])? DMAT : DMIS;
        else
            mis += _day[*px-'A'][*py-'A'];

10      /* update penalty for del in x seq;
        * favor new del over ongong del
        * ignore MAXGAP if weighting endgaps
        */
        if (endgaps || ndely[yy] < MAXGAP) {
15            if (col0[yy] - ins0 >= dely[yy]) {
                dely[yy] = col0[yy] - (ins0+ins1);
                ndely[yy] = 1;
            } else {
                dely[yy] -= ins1;
                ndely[yy]++;
20            }
        } else {
            if (col0[yy] - (ins0+ins1) >= dely[yy]) {
                dely[yy] = col0[yy] - (ins0+ins1);
                ndely[yy] = 1;
25            } else
                ndely[yy]++;
        }

        /* update penalty for del in y seq;
        * favor new del over ongong del
        */
        if (endgaps || ndelx < MAXGAP) {
30            if (col1[yy-1] - ins0 >= delx) {
                delx = col1[yy-1] - (ins0+ins1);
                ndelx = 1;
35            } else {
                delx -= ins1;
                ndelx++;
            }
        } else {
40            if (col1[yy-1] - (ins0+ins1) >= delx) {
                delx = col1[yy-1] - (ins0+ins1);
                ndelx = 1;
            } else
45                ndelx++;
        }

        /* pick the maximum score; we're favoring
50        * mis over any del and delx over dely
        */
55
60

```

**Table 1 (cont')**

...nw

```

0      id = xx - yy + len1 - 1;
      if (mis >= delx && mis >= dely[yy])
          coll[yy] = mis;
5      else if (delx >= dely[yy]) {
          coll[yy] = delx;
          ij = dx[id].ijmp;
          if (dx[id].jp.n[0] && (!dna || (ndelx >= MAXJMP
10      && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
              dx[id].ijmp++;
              if (++ij >= MAXJMP) {
                  writejumps(id);
                  ij = dx[id].ijmp = 0;
                  dx[id].offset = offset;
15      offset += sizeof(struct jmp) + sizeof(offset);
              }
          }
          dx[id].jp.n[ij] = ndelx;
          dx[id].jp.x[ij] = xx;
          dx[id].score = delx;
20      }
      else {
          coll[yy] = dely[yy];
          ij = dx[id].ijmp;
25      if (dx[id].jp.n[0] && (!dna || (ndely[yy] >= MAXJMP
          && xx > dx[id].jp.x[ij]+MX) || mis > dx[id].score+DINS0)) {
              dx[id].ijmp++;
              if (++ij >= MAXJMP) {
                  writejumps(id);
                  ij = dx[id].ijmp = 0;
                  dx[id].offset = offset;
30      offset += sizeof(struct jmp) + sizeof(offset);
              }
          }
          dx[id].jp.n[ij] = -ndely[yy];
          dx[id].jp.x[ij] = xx;
          dx[id].score = dely[yy];
35      }
      if (xx == len0 && yy < len1) {
          /* last col
          */
          if (endgaps)
              coll[yy] -= ins0 + ins1*(len1-yy);
          if (coll[yy] > smax) {
45      smax = coll[yy];
              dmax = id;
          }
      }
  }
50      if (endgaps && xx < len0)
          coll[yy-1] -= ins0 + ins1*(len0-xx);
      if (coll[yy-1] > smax) {
          smax = coll[yy-1];
          dmax = id;
55      }
      tmp = col0; col0 = coll; coll = tmp;
  }
  (void) free((char *)ndely);
  (void) free((char *)dely);
60  (void) free((char *)col0);
  (void) free((char *)coll);
  }

```

**Table 1 (cont')**

```

/*
 *
 * print() -- only routine visible outside this module
 *
5  * static:
 * getmat() -- trace back best path, count matches: print()
 * pr_align() -- print alignment of described in array p[]: print()
 * dumpblock() -- dump a block of lines with numbers, stars: pr_align()
 * nums() -- put out a number line: dumpblock()
10 * putline() -- put out a line (name, [num], seq, [num]): dumpblock()
 * stars() -- put a line of stars: dumpblock()
 * stripname() -- strip any path and prefix from a seqname
 */

15 #include "nw.h"

#define SPC      3
#define P_LINE  256 /* maximum output line */
#define P_SPC    3   /* space between name or num and seq */

20 extern _day[26][26];
int olen; /* set output line length */
FILE *fx; /* output file */

25 print()
{
    int lx, ly, firstgap, lastgap; /* overlap */

    if ((fx = fopen(ofile, "w")) == 0) {
30         fprintf(stderr, "%s: can't write %s\n", prog, ofile);
        cleanup(1);
    }
    fprintf(fx, "< first sequence: %s (length = %d)\n", namex[0], len0);
    fprintf(fx, "< second sequence: %s (length = %d)\n", namex[1], len1);
35     olen = 60;
    lx = len0;
    ly = len1;
    firstgap = lastgap = 0;
    if (dmax < len1 - 1) { /* leading gap in x */
40         pp[0].spc = firstgap = len1 - dmax - 1;
        ly -= pp[0].spc;
    }
    else if (dmax > len1 - 1) { /* leading gap in y */
45         pp[1].spc = firstgap = dmax - (len1 - 1);
        lx -= pp[1].spc;
    }
    if (dmax0 < len0 - 1) { /* trailing gap in x */
50         lastgap = len0 - dmax0 - 1;
        lx -= lastgap;
    }
    else if (dmax0 > len0 - 1) { /* trailing gap in y */
55         lastgap = dmax0 - (len0 - 1);
        ly -= lastgap;
    }
    getmat(lx, ly, firstgap, lastgap);
    pr_align();
}

60

```

**print**

Table 1 (cont')

```

/*
 * trace back the best path, count matches
 */
static
5  getmat(lx, ly, firstgap, lastgap)                                getmat
    int      lx, ly;                                /* "core" (minus endgaps) */
    int      firstgap, lastgap;                      /* leading trailing overlap */
{
    int      nm, i0, i1, siz0, siz1;
10    char     outx[32];
    double   pct;
    register n0, n1;
    register char *p0, *p1;

15    /* get total matches, score
    */
    i0 = i1 = siz0 = siz1 = 0;
    p0 = seqx[0] + pp[1].spc;
    p1 = seqx[1] + pp[0].spc;
20    n0 = pp[1].spc + 1;
    n1 = pp[0].spc + 1;

    nm = 0;
    while ( *p0 && *p1 ) {
25         if (siz0) {
            p1++;
            n1++;
            siz0--;
        }
30         else if (siz1) {
            p0++;
            n0++;
            siz1--;
        }
35         else {
            if (xbm[*p0-'A']&xbm[*p1-'A'])
                nm++;
            if (n0++ == pp[0].x[i0])
                siz0 = pp[0].n[i0++];
40             if (n1++ == pp[1].x[i1])
                siz1 = pp[1].n[i1++];
            p0++;
            p1++;
        }
45     }

    /* pct homology:
    * if penalizing endgaps, base is the shorter seq
    * else, knock off overhangs and take shorter core
    */
50     if (endgaps)
        lx = (len0 < len1)? len0 : len1;
    else
        lx = (lx < ly)? lx : ly;
55     pct = 100.*(double)nm/(double)lx;
    fprintf(fx, "\n");
    fprintf(fx, "< %d match%s in an overlap of %d: %.2f percent similarity\n",
        nm, (nm == 1)? "" : "es", lx, pct);
60

```

**Table 1 (cont')**

```

fprintf(fx, "< gaps in first sequence: %d", gapx);
if (gapx) {
    (void) sprintf(outx, " (%d %s%s)",
        ngapx, (dna)? "base":"residue", (ngapx == 1)? "" : "s");
    fprintf(fx, "%s", outx);

    fprintf(fx, ", gaps in second sequence: %d", gapy);
    if (gapy) {
        (void) sprintf(outx, " (%d %s%s)",
            ngapy, (dna)? "base":"residue", (ngapy == 1)? "" : "s");
        fprintf(fx, "%s", outx);
    }
    if (dna)
        fprintf(fx,
            "\n<score: %d (match = %d, mismatch = %d, gap penalty = %d + %d per base)\n",
            smax, DMAT, DMIS, DINS0, DINS1);
    else
        fprintf(fx,
            "\n<score: %d (Dayhoff PAM 250 matrix, gap penalty = %d + %d per residue)\n",
            smax, PINS0, PINS1);
    if (endgaps)
        fprintf(fx,
            "<endgaps penalized. left endgap: %d %s%s, right endgap: %d %s%s\n",
            firstgap, (dna)? "base" : "residue", (firstgap == 1)? "" : "s",
            lastgap, (dna)? "base" : "residue", (lastgap == 1)? "" : "s");
    else
        fprintf(fx, "<endgaps not penalized\n");
}

static      nm;          /* matches in core -- for checking */
static      lmax;        /* lengths of stripped file names */
static      ij[2];       /* jmp index for a path */
static      nc[2];       /* number at start of current line */
static      ni[2];       /* current elem number -- for gapping */
static      siz[2];
static char  *ps[2];      /* ptr to current element */
static char  *po[2];      /* ptr to next output char slot */
static char  out[2][P_LINE]; /* output line */
static char  star[P_LINE]; /* set by stars() */

/*
 * print alignment of described in struct path pp[]
 */
static
pr_align()
{
    int      nn;          /* char count */
    int      more;
    register i;

    for (i = 0, lmax = 0; i < 2; i++) {
        nn = stripname(namex[i]);
        if (nn > lmax)
            lmax = nn;

        nc[i] = 1;
        ni[i] = 1;
        siz[i] = ij[i] = 0;
        ps[i] = seqx[i];
        po[i] = out[i];
    }
}

```

...getmat

pr\_align



**Table 1 (cont')**

```

for (nn = nm = 0, more = 1; more; ) {
    for (i = more = 0; i < 2; i++) {
        /*
5         * do we have more of this sequence?
        */
        if (!*ps[i])
            continue;

10         more++;

        if (pp[i].spc) { /* leading space */
            *po[i]++ = ' ';
            pp[i].spc--;
15         }
        else if (siz[i]) { /* in a gap */
            *po[i]++ = '-';
            siz[i]--;
20         }
        else { /* we're putting a seq element
            */
            *po[i] = *ps[i];
            if (islower(*ps[i]))
                *ps[i] = toupper(*ps[i]);
25             po[i]++;
            ps[i]++;

            /*
30             * are we at next gap for this seq?
            */
            if (ni[i] == pp[i].x[ij[i]]) {
                /*
                * we need to merge all gaps
                * at this location
35                 */
                siz[i] = pp[i].n[ij[i] + +];
                while (ni[i] == pp[i].x[ij[i]])
                    siz[i] += pp[i].n[ij[i] + +];

40                 }
                ni[i]++;
            }
        }
        if (++nn == olen || !more && nn) {
45             dumpblock();
            for (i = 0; i < 2; i++)
                po[i] = out[i];
            nn = 0;
        }
50     }
}

/*
 * dump a block of lines, including numbers, stars: pr_align()
 */
55 static
dumpblock()
{
    register i;

60     for (i = 0; i < 2; i++)
        *po[i] = '\0';

```

**dumpblock**

Table 1 (cont')

...dumpblock

```

5      (void)putc('\n', fx);
      for (i = 0; i < 2; i++) {
          if (*out[i] && (*out[i] != ' ' || *(po[i]) != ' ')) {
              if (i == 0)
                  nums(i);
              if (i == 0 && *out[1])
                  stars();
10             putline(i);
              if (i == 0 && *out[1])
                  fprintf(fx, star);
              if (i == 1)
                  nums(i);
15             }
          }
      }

20  /*
      * put out a number line: dumpblock()
      */
      static
      nums(ix)
25  {
          int      ix;      /* index in out[] holding seq line */

          char      nline[P_LINE];
          register  i, j;
          register char *pn, *px, *py;

30      for (pn = nline, i = 0; i < lmax+P_SPC; i++, pn++)
          *pn = ' ';
          for (i = nc[ix], py = out[ix]; *py; py++, pn++) {
              if (*py == ' ' || *py == '-')
                  *pn = ' ';
35              else {
                  if (i%10 == 0 || (i == 1 && nc[ix] != 1)) {
                      j = (i < 0)? -i : i;
                      for (px = pn; j /= 10, px--)
                          *px = j%10 + '0';
40                      if (i < 0)
                          *px = '-';

                      }
                  else
                      *pn = ' ';
45                      i++;
                  }
              }
          *pn = '\0';
          nc[ix] = i;
50      for (pn = nline; *pn; pn++)
          (void)putc(*pn, fx);
          (void)putc('\n', fx);
      }

55  /*
      * put out a line (name, [num], seq, [num]): dumpblock()
      */
      static
      putline(ix)
60      int      ix;      {

```

nums

putline

**Table 1 (cont')**

...putline

```

5      int          i;
      register char *px;

      for (px = namex[ix], i = 0; *px && *px != ':'; px++, i++)
          (void) putc(*px, fx);
      for (; i < lmax+P_SPC; i++)
          (void) putc(' ', fx);

10     /* these count from 1:
       * ni[] is current element (from 1)
       * nc[] is number at start of current line
       */
15     for (px = out[ix]; *px; px++)
          (void) putc(*px&0x7F, fx);
      (void) putc('\n', fx);
  }

20  /*
   * put a line of stars (seqs always in out[0], out[1]): dumpblock()
   */
   static
25  stars()
  {
      int          i;
      register char *p0, *p1, cx, *px;

30     if (!*out[0] || (*out[0] == ' ' && *(po[0]) == ' ') ||
        !*out[1] || (*out[1] == ' ' && *(po[1]) == ' '))
          return;
      px = star;
      for (i = lmax+P_SPC; i; i--)
35         *px++ = ' ';

      for (p0 = out[0], p1 = out[1]; *p0 && *p1; p0++, p1++) {
          if (isalpha(*p0) && isalpha(*p1)) {
40             if (xbm[*p0-'A']&xbm[*p1-'A']) {
                 cx = '*';
                 nm++;
             }
             else if (!dna && _day[*p0-'A'][*p1-'A'] > 0)
45                 cx = '.';
             else
                 cx = ' ';
          }
          else
50             cx = ' ';
          *px++ = cx;
      }
      *px++ = '\n';
      *px = '\0';
55  }

```

stars

60

**Table 1 (cont')**

```

/*
 * strip path or prefix from pn, return len: pr_align()
 */
static
5 stripname(pn)                                stripname
{
    char    *pn;    /* file name (may be path) */
    register char    *px, *py;
10     py = 0;
    for (px = pn; *px; px++)
        if (*px == '/')
            py = px + 1;
15     if (py)
        (void) strcpy(pn, py);
    return(strlen(pn));
}
20
25
30
35
40
45
50
55
60

```

**Table 1 (cont')**

```

/*
 * cleanup() -- cleanup any tmp file
 * getseq() -- read in seq, set dna, len, maxlen
 * g_calloc() -- calloc() with error checkin
5  * readjumps() -- get the good jumps; from tmp file if necessary
 * writejumps() -- write a filled array of jumps to a tmp file: nw()
 */
#include "nw.h"
#include <sys/file.h>

10 char *jname = "/tmp/homgXXXXXX"; /* tmp file for jumps */
FILE *fj;

int cleanup(); /* cleanup tmp file */
15 long lseek();

/*
 * remove any tmp file if we blow
 */
20 cleanup(i)
    int i;
{
    if (fj)
        (void) unlink(jname);
25    exit(i);
}

/*
 * read, return ptr to seq, set dna, len, maxlen
 * skip lines starting with ';', '<', or '>'
 * seq in upper or lower case
 */
30 char *
getseq(file, len)
35 char *file; /* file name */
    int *len; /* seq len */
{
    char line[1024], *pseq;
    register char *px, *py;
    int natgc, tlen;
    FILE *fp;

    if ((fp = fopen(file, "r")) == 0) {
        fprintf(stderr, "%s: can't read %s\n", prog, file);
        exit(1);
    }
    tlen = natgc = 0;
    while (fgets(line, 1024, fp)) {
        if (*line == ';' || *line == '<' || *line == '>')
            continue;
        for (px = line; *px != '\n'; px++)
            if (isupper(*px) || islower(*px))
                tlen++;
    }
    if ((pseq = malloc((unsigned)(tlen+6))) == 0) {
        fprintf(stderr, "%s: malloc() failed to get %d bytes for %s\n", prog, tlen+6, file);
        exit(1);
    }
    pseq[0] = pseq[1] = pseq[2] = pseq[3] = '\0';
60

```

**Table 1 (cont')**

...getseq

```

py = pseq + 4;
*len = tlen;
rewind(fp);
5
while (fgets(line, 1024, fp)) {
    if (*line == ';' || *line == '<' || *line == '>')
        continue;
    for (px = line; *px != '\n'; px++) {
10        if (isupper(*px))
            *py++ = *px;
        else if (islower(*px))
            *py++ = toupper(*px);
        if (index("ATGCU", *(py-1)))
15            natgc++;
    }
    *py++ = '\0';
    *py = '\0';
20    (void) fclose(fp);
    dna = natgc > (tlen/3);
    return(pseq+4);
}

25 char *
g_alloc(msg, nx, sz)
char *msg; /* program, calling routine */
int nx, sz; /* number and size of elements */
{
30     char *px, *calloc();

    if ((px = calloc((unsigned)nx, (unsigned)sz)) == 0) {
        if (*msg) {
35             fprintf(stderr, "%s: g_alloc() failed %s (n=%d, sz=%d)\n", prog, msg, nx, sz);
            exit(1);
        }
    }
    return(px);
40 }

/*
 * get final jmps from dx[] or tmp file, set pp[], reset dmax: main()
 */
readjmps()
45 {
    int fd = -1;
    int siz, i0, i1;
    register i, j, xx;

50     if (fj) {
        (void) fclose(fj);
        if ((fd = open(jname, O_RDONLY, 0)) < 0) {
            fprintf(stderr, "%s: can't open() %s\n", prog, jname);
            cleanup(1);
55         }
    }
    for (i = i0 = i1 = 0, dmax0 = dmax, xx = len0; i++) {
        while (1) {
60             for (j = dx[dmax].ijmp; j >= 0 && dx[dmax].jp.x[j] >= xx; j--)
                ;

```

g\_alloc

readjmps

**Table 1 (cont')****...readjumps**

```

5      if (j < 0 && dx[dmax].offset && fj) {
        (void) lseek(fd, dx[dmax].offset, 0);
        (void) read(fd, (char *)&dx[dmax].jp, sizeof(struct jmp));
        (void) read(fd, (char *)&dx[dmax].offset, sizeof(dx[dmax].offset));
        dx[dmax].ijmp = MAXJMP-1;
      }
      else
10         break;
    }
    if (i >= JMPS) {
        fprintf(stderr, "%s: too many gaps in alignment\n", prog);
        cleanup(1);
    }
15    if (j >= 0) {
        siz = dx[dmax].jp.n[j];
        xx = dx[dmax].jp.x[j];
        dmax += siz;
        if (siz < 0) { /* gap in second seq */
20            pp[1].n[i1] = -siz;
            xx += siz;
            /* id = xx - yy + len1 - 1
             */
            pp[1].x[i1] = xx - dmax + len1 - 1;
            gapy++;
            ngapy -= siz;
            /* ignore MAXGAP when doing endgaps */
            siz = (-siz < MAXGAP || endgaps)? -siz : MAXGAP;
            i1++;
30        }
        else if (siz > 0) { /* gap in first seq */
            pp[0].n[i0] = siz;
            pp[0].x[i0] = xx;
            gapx++;
            ngapx += siz;
35            /* ignore MAXGAP when doing endgaps */
            siz = (siz < MAXGAP || endgaps)? siz : MAXGAP;
            i0++;
        }
40    }
    else
        break;
}

45    /* reverse the order of jumps
    */
    for (j = 0, i0--; j < i0; j++, i0--) {
        i = pp[0].n[j]; pp[0].n[j] = pp[0].n[i0]; pp[0].n[i0] = i;
        i = pp[0].x[j]; pp[0].x[j] = pp[0].x[i0]; pp[0].x[i0] = i;
50    }
    for (j = 0, i1--; j < i1; j++, i1--) {
        i = pp[1].n[j]; pp[1].n[j] = pp[1].n[i1]; pp[1].n[i1] = i;
        i = pp[1].x[j]; pp[1].x[j] = pp[1].x[i1]; pp[1].x[i1] = i;
55    }
    if (fd >= 0)
        (void) close(fd);
    if (fj) {
        (void) unlink(jname);
        fj = 0;
60        offset = 0;
    }
}

```

**Table 1 (cont')**

```

/*
 * write a filled jmp struct offset of the prev one (if any): nw()
 */
5  writejumps(ix)                                writejumps
    int    ix;
    {
        char    *mktemp();
10         if (!fj) {
            if (mktemp(jname) < 0) {
                fprintf(stderr, "%s: can't mktemp() %s\n", prog, jname);
                cleanup(1);
            }
15             if ((fj = fopen(jname, "w")) == 0) {
                fprintf(stderr, "%s: can't write %s\n", prog, jname);
                exit(1);
            }
20         }
        (void) fwrite((char *)&dx[ix].jp, sizeof(struct jmp), 1, fj);
        (void) fwrite((char *)&dx[ix].offset, sizeof(dx[ix].offset), 1, fj);
    }
}
25
30
35
40
45
50
55
60

```



**Table 2**

PRO	XXXXXXXXXXXXXXXXXX	(Length = 15 amino acids)
Comparison Protein	XXXXXXXXYYYYYY	(Length = 12 amino acids)

5    % amino acid sequence identity =

(the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

10    5 divided by 15 = 33.3%

**Table 3**

PRO	XXXXXXXXXX	(Length = 10 amino acids)
15    Comparison Protein	XXXXXXXXYYYYYYZZYZ	(Length = 15 amino acids)

% amino acid sequence identity =

20    (the number of identically matching amino acid residues between the two polypeptide sequences as determined by ALIGN-2) divided by (the total number of amino acid residues of the PRO polypeptide) =

5 divided by 10 = 50%

**Table 4**

25    PRO-DNA	NNNNNNNNNNNNNN	(Length = 14 nucleotides)
Comparison DNA	NNNNNNLLLLLLLLLL	(Length = 16 nucleotides)

% nucleic acid sequence identity =

30

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

6 divided by 14 = 42.9%

35

Table 5

PRO-DNA	NNNNNNNNNNNN	(Length = 12 nucleotides)
Comparison DNA	NNNNLLLTVV	(Length = 9 nucleotides)

5     % nucleic acid sequence identity =

(the number of identically matching nucleotides between the two nucleic acid sequences as determined by ALIGN-2) divided by (the total number of nucleotides of the PRO-DNA nucleic acid sequence) =

10    4 divided by 12 = 33.3%

## II.           Compositions and Methods of the Invention

### A.       Full-Length PRO Polypeptides

15     The present invention provides newly identified and isolated nucleotide sequences encoding polypeptides referred to in the present application as PRO polypeptides. In particular, cDNAs encoding various PRO polypeptides have been identified and isolated, as disclosed in further detail in the Examples below. It is noted that proteins produced in separate expression rounds may be given different PRO numbers but the UNQ number is unique for any given DNA and the encoded protein, and will not be changed. However, for sake of simplicity, in the present specification the protein encoded by the full length native nucleic acid molecules disclosed herein as well as all further native homologues and variants included in the foregoing definition of PRO, will be referred to as "PRO/number", regardless of their origin or mode of preparation.

20     As disclosed in the Examples below, various cDNA clones have been deposited with the ATCC. The actual nucleotide sequences of those clones can readily be determined by the skilled artisan by sequencing of the deposited clone using routine methods in the art. The predicted amino acid sequence can be determined from the nucleotide sequence using routine skill. For the PRO polypeptides and encoding nucleic acids described herein, Applicants have identified what is believed to be the reading frame best identifiable with the sequence information available at the time.

### B.       PRO Polypeptide Variants

30     In addition to the full-length native sequence PRO polypeptides described herein, it is contemplated that PRO variants can be prepared. PRO variants can be prepared by introducing appropriate nucleotide changes into the PRO DNA, and/or by synthesis of the desired PRO polypeptide. Those skilled in the art will appreciate that amino acid changes may alter post-translational processes of the PRO, such as changing the number or position of glycosylation sites or altering the membrane anchoring characteristics.

35     Variations in the native full-length sequence PRO or in various domains of the PRO described herein, can be made, for example, using any of the techniques and guidelines for conservative and non-conservative

mutations set forth, for instance, in U.S. Patent No. 5,364,934. Variations may be a substitution, deletion or insertion of one or more codons encoding the PRO that results in a change in the amino acid sequence of the PRO as compared with the native sequence PRO. Optionally the variation is by substitution of at least one amino acid with any other amino acid in one or more of the domains of the PRO. Guidance in determining which amino acid residue may be inserted, substituted or deleted without adversely affecting the desired activity may be found by  
5 comparing the sequence of the PRO with that of homologous known protein molecules and minimizing the number of amino acid sequence changes made in regions of high homology. Amino acid substitutions can be the result of replacing one amino acid with another amino acid having similar structural and/or chemical properties, such as the replacement of a leucine with a serine, i.e., conservative amino acid replacements. Insertions or deletions may optionally be in the range of about 1 to 5 amino acids. The variation allowed may be determined by  
10 systematically making insertions, deletions or substitutions of amino acids in the sequence and testing the resulting variants for activity exhibited by the full-length or mature native sequence.

PRO polypeptide fragments are provided herein. Such fragments may be truncated at the N-terminus or C-terminus, or may lack internal residues, for example, when compared with a full length native protein. Certain fragments lack amino acid residues that are not essential for a desired biological activity of the PRO  
15 polypeptide.

PRO fragments may be prepared by any of a number of conventional techniques. Desired peptide fragments may be chemically synthesized. An alternative approach involves generating PRO fragments by enzymatic digestion, e.g., by treating the protein with an enzyme known to cleave proteins at sites defined by particular amino acid residues, or by digesting the DNA with suitable restriction enzymes and isolating the desired  
20 fragment. Yet another suitable technique involves isolating and amplifying a DNA fragment encoding a desired polypeptide fragment, by polymerase chain reaction (PCR). Oligonucleotides that define the desired termini of the DNA fragment are employed at the 5' and 3' primers in the PCR. Preferably, PRO polypeptide fragments share at least one biological and/or immunological activity with the native PRO polypeptide disclosed herein.

In particular embodiments, conservative substitutions of interest are shown in Table 6 under the heading  
25 of preferred substitutions. If such substitutions result in a change in biological activity, then more substantial changes, denominated exemplary substitutions in Table 6, or as further described below in reference to amino acid classes, are introduced and the products screened.

Table 6

	<u>Original Residue</u>	<u>Exemplary Substitutions</u>	<u>Preferred Substitutions</u>
5	Ala (A)	val; leu; ile	val
	Arg (R)	lys; gln; asn	lys
	Asn (N)	gln; his; lys; arg	gln
	Asp (D)	glu	glu
	Cys (C)	ser	ser
10	Gln (Q)	asn	asn
	Glu (E)	asp	asp
	Gly (G)	pro; ala	ala
	His (H)	asn; gln; lys; arg	arg
	Ile (I)	leu; val; met; ala; phe;	
15		norleucine	leu
	Leu (L)	norleucine; ile; val;	
		met; ala; phe	ile
	Lys (K)	arg; gln; asn	arg
	Met (M)	leu; phe; ile	leu
20	Phe (F)	leu; val; ile; ala; tyr	leu
	Pro (P)	ala	ala
	Ser (S)	thr	thr
	Thr (T)	ser	ser
	Trp (W)	tyr; phe	tyr
25	Tyr (Y)	trp; phe; thr; ser	phe
	Val (V)	ile; leu; met; phe;	
		ala; norleucine	leu

30 Substantial modifications in function or immunological identity of the PRO polypeptide are accomplished by selecting substitutions that differ significantly in their effect on maintaining (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a sheet or helical conformation, (b) the charge or hydrophobicity of the molecule at the target site, or (c) the bulk of the side chain. Naturally occurring residues are divided into groups based on common side-chain properties:

- 35 (1) hydrophobic: norleucine, met, ala, val, leu, ile;  
 (2) neutral hydrophilic: cys, ser, thr;  
 (3) acidic: asp, glu;  
 (4) basic: asn, gln, his, lys, arg;  
 (5) residues that influence chain orientation: gly, pro; and  
 40 (6) aromatic: trp, tyr, phe.

Non-conservative substitutions will entail exchanging a member of one of these classes for another class. Such substituted residues also may be introduced into the conservative substitution sites or, more preferably, into the remaining (non-conserved) sites.

45 The variations can be made using methods known in the art such as oligonucleotide-mediated (site-directed) mutagenesis, alanine scanning, and PCR mutagenesis. Site-directed mutagenesis [Carter et al., Nucl. Acids Res., 13:4331 (1986); Zoller et al., Nucl. Acids Res., 10:6487 (1987)], cassette mutagenesis [Wells et al.,

Gene, 34:315 (1985)], restriction selection mutagenesis [Wells et al., Philos. Trans. R. Soc. London SerA, 317:415 (1986)] or other known techniques can be performed on the cloned DNA to produce the PRO variant DNA.

Scanning amino acid analysis can also be employed to identify one or more amino acids along a contiguous sequence. Among the preferred scanning amino acids are relatively small, neutral amino acids. Such amino acids include alanine, glycine, serine, and cysteine. Alanine is typically a preferred scanning amino acid among this group because it eliminates the side-chain beyond the beta-carbon and is less likely to alter the main-chain conformation of the variant [Cunningham and Wells, Science, 244: 1081-1085 (1989)]. Alanine is also typically preferred because it is the most common amino acid. Further, it is frequently found in both buried and exposed positions [Creighton, The Proteins, (W.H. Freeman & Co., N.Y.); Chothia, J. Mol. Biol., 150:1 (1976)]. If alanine substitution does not yield adequate amounts of variant, an isoteric amino acid can be used.

### C. Modifications of PRO

Covalent modifications of PRO are included within the scope of this invention. One type of covalent modification includes reacting targeted amino acid residues of a PRO polypeptide with an organic derivatizing agent that is capable of reacting with selected side chains or the N- or C- terminal residues of the PRO. Derivatization with bifunctional agents is useful, for instance, for crosslinking PRO to a water-insoluble support matrix or surface for use in the method for purifying anti-PRO antibodies, and vice-versa. Commonly used crosslinking agents include, e.g., 1,1-bis(diazoacetyl)-2-phenylethane, glutaraldehyde, N-hydroxysuccinimide esters, for example, esters with 4-azidosalicylic acid, homobifunctional imidoesters, including disuccinimidyl esters such as 3,3'-dithiobis(succinimidylpropionate), bifunctional maleimides such as bis-N-maleimido-1,8-octane and agents such as methyl-3-[(p-azidophenyl)dithio]propioimide.

Other modifications include deamidation of glutaminyl and asparaginy residues to the corresponding glutamyl and aspartyl residues, respectively, hydroxylation of proline and lysine, phosphorylation of hydroxyl groups of seryl or threonyl residues, methylation of the  $\alpha$ -amino groups of lysine, arginine, and histidine side chains [T.E. Creighton, Proteins: Structure and Molecular Properties, W.H. Freeman & Co., San Francisco, pp. 79-86 (1983)], acetylation of the N-terminal amine, and amidation of any C-terminal carboxyl group.

Another type of covalent modification of the PRO polypeptide included within the scope of this invention comprises altering the native glycosylation pattern of the polypeptide. "Altering the native glycosylation pattern" is intended for purposes herein to mean deleting one or more carbohydrate moieties found in native sequence PRO (either by removing the underlying glycosylation site or by deleting the glycosylation by chemical and/or enzymatic means), and/or adding one or more glycosylation sites that are not present in the native sequence PRO. In addition, the phrase includes qualitative changes in the glycosylation of the native proteins, involving a change in the nature and proportions of the various carbohydrate moieties present.

Addition of glycosylation sites to the PRO polypeptide may be accomplished by altering the amino acid sequence. The alteration may be made, for example, by the addition of, or substitution by, one or more serine or threonine residues to the native sequence PRO (for O-linked glycosylation sites). The PRO amino acid sequence may optionally be altered through changes at the DNA level, particularly by mutating the DNA encoding

the PRO polypeptide at preselected bases such that codons are generated that will translate into the desired amino acids.

Another means of increasing the number of carbohydrate moieties on the PRO polypeptide is by chemical or enzymatic coupling of glycosides to the polypeptide. Such methods are described in the art, e.g., in WO 87/05330 published 11 September 1987, and in Aplin and Wriston, CRC Crit. Rev. Biochem., pp. 259-306 (1981).

Removal of carbohydrate moieties present on the PRO polypeptide may be accomplished chemically or enzymatically or by mutational substitution of codons encoding for amino acid residues that serve as targets for glycosylation. Chemical deglycosylation techniques are known in the art and described, for instance, by Hakimuddin, et al., Arch. Biochem. Biophys., 259:52 (1987) and by Edge et al., Anal. Biochem., 118:131 (1981). Enzymatic cleavage of carbohydrate moieties on polypeptides can be achieved by the use of a variety of endo- and exo-glycosidases as described by Thotakura et al., Meth. Enzymol., 138:350 (1987).

Another type of covalent modification of PRO comprises linking the PRO polypeptide to one of a variety of nonproteinaceous polymers, e.g., polyethylene glycol (PEG), polypropylene glycol, or polyoxyalkylenes, in the manner set forth in U.S. Patent Nos. 4,640,835; 4,496,689; 4,301,144; 4,670,417; 4,791,192 or 4,179,337.

The PRO of the present invention may also be modified in a way to form a chimeric molecule comprising PRO fused to another, heterologous polypeptide or amino acid sequence.

In one embodiment, such a chimeric molecule comprises a fusion of the PRO with a tag polypeptide which provides an epitope to which an anti-tag antibody can selectively bind. The epitope tag is generally placed at the amino- or carboxyl- terminus of the PRO. The presence of such epitope-tagged forms of the PRO can be detected using an antibody against the tag polypeptide. Also, provision of the epitope tag enables the PRO to be readily purified by affinity purification using an anti-tag antibody or another type of affinity matrix that binds to the epitope tag. Various tag polypeptides and their respective antibodies are well known in the art. Examples include poly-histidine (poly-his) or poly-histidine-glycine (poly-his-gly) tags; the flu HA tag polypeptide and its antibody 12CA5 [Field et al., Mol. Cell. Biol., 8:2159-2165 (1988)]; the c-myc tag and the 8F9, 3C7, 6E10, G4, B7 and 9E10 antibodies thereto [Evan et al., Molecular and Cellular Biology, 5:3610-3616 (1985)]; and the Herpes Simplex virus glycoprotein D (gD) tag and its antibody [Paborsky et al., Protein Engineering, 3(6):547-553 (1990)]. Other tag polypeptides include the Flag-peptide [Hopp et al., BioTechnology, 6:1204-1210 (1988)]; the KT3 epitope peptide [Martin et al., Science, 255:192-194 (1992)]; an  $\alpha$ -tubulin epitope peptide [Skinner et al., J. Biol. Chem., 266:15163-15166 (1991)]; and the T7 gene 10 protein peptide tag [Lutz-Freyermuth et al., Proc. Natl. Acad. Sci. USA, 87:6393-6397 (1990)].

In an alternative embodiment, the chimeric molecule may comprise a fusion of the PRO with an immunoglobulin or a particular region of an immunoglobulin. For a bivalent form of the chimeric molecule (also referred to as an "immunoadhesin"), such a fusion could be to the Fc region of an IgG molecule. The Ig fusions preferably include the substitution of a soluble (transmembrane domain deleted or inactivated) form of a PRO polypeptide in place of at least one variable region within an Ig molecule. In a particularly preferred embodiment, the immunoglobulin fusion includes the hinge, CH2 and CH3, or the hinge, CH1, CH2 and CH3 regions of an IgG1 molecule. For the production of immunoglobulin fusions see also US Patent No. 5,428,130 issued June 27,

1995.

#### D. Preparation of PRO

The description below relates primarily to production of PRO by culturing cells transformed or transfected with a vector containing PRO nucleic acid. It is, of course, contemplated that alternative methods, which are well known in the art, may be employed to prepare PRO. For instance, the PRO sequence, or portions thereof, may be produced by direct peptide synthesis using solid-phase techniques [see, e.g., Stewart et al., Solid-Phase Peptide Synthesis, W.H. Freeman Co., San Francisco, CA (1969); Merrifield, J. Am. Chem. Soc., 85:2149-2154 (1963)]. *In vitro* protein synthesis may be performed using manual techniques or by automation. Automated synthesis may be accomplished, for instance, using an Applied Biosystems Peptide Synthesizer (Foster City, CA) using manufacturer's instructions. Various portions of the PRO may be chemically synthesized separately and combined using chemical or enzymatic methods to produce the full-length PRO.

##### 1. Isolation of DNA Encoding PRO

DNA encoding PRO may be obtained from a cDNA library prepared from tissue believed to possess the PRO mRNA and to express it at a detectable level. Accordingly, human PRO DNA can be conveniently obtained from a cDNA library prepared from human tissue, such as described in the Examples. The PRO-encoding gene may also be obtained from a genomic library or by known synthetic procedures (e.g., automated nucleic acid synthesis).

Libraries can be screened with probes (such as antibodies to the PRO or oligonucleotides of at least about 20-80 bases) designed to identify the gene of interest or the protein encoded by it. Screening the cDNA or genomic library with the selected probe may be conducted using standard procedures, such as described in Sambrook et al., Molecular Cloning: A Laboratory Manual (New York: Cold Spring Harbor Laboratory Press, 1989). An alternative means to isolate the gene encoding PRO is to use PCR methodology [Sambrook et al., supra; Dieffenbach et al., PCR Primer: A Laboratory Manual (Cold Spring Harbor Laboratory Press, 1995)].

The Examples below describe techniques for screening a cDNA library. The oligonucleotide sequences selected as probes should be of sufficient length and sufficiently unambiguous that false positives are minimized. The oligonucleotide is preferably labeled such that it can be detected upon hybridization to DNA in the library being screened. Methods of labeling are well known in the art, and include the use of radiolabels like <sup>32</sup>P-labeled ATP, biotinylation or enzyme labeling. Hybridization conditions, including moderate stringency and high stringency, are provided in Sambrook et al., supra.

Sequences identified in such library screening methods can be compared and aligned to other known sequences deposited and available in public databases such as GenBank or other private sequence databases. Sequence identity (at either the amino acid or nucleotide level) within defined regions of the molecule or across the full-length sequence can be determined using methods known in the art and as described herein.

Nucleic acid having protein coding sequence may be obtained by screening selected cDNA or genomic libraries using the deduced amino acid sequence disclosed herein for the first time, and, if necessary, using conventional primer extension procedures as described in Sambrook et al., supra, to detect precursors and

processing intermediates of mRNA that may not have been reverse-transcribed into cDNA.

## 2. Selection and Transformation of Host Cells

Host cells are transfected or transformed with expression or cloning vectors described herein for PRO production and cultured in conventional nutrient media modified as appropriate for inducing promoters, selecting transformants, or amplifying the genes encoding the desired sequences. The culture conditions, such as media, temperature, pH and the like, can be selected by the skilled artisan without undue experimentation. In general, principles, protocols, and practical techniques for maximizing the productivity of cell cultures can be found in Mammalian Cell Biotechnology: a Practical Approach, M. Butler, ed. (IRL Press, 1991) and Sambrook et al., supra.

Methods of eukaryotic cell transfection and prokaryotic cell transformation are known to the ordinarily skilled artisan, for example,  $\text{CaCl}_2$ ,  $\text{CaPO}_4$ , liposome-mediated and electroporation. Depending on the host cell used, transformation is performed using standard techniques appropriate to such cells. The calcium treatment employing calcium chloride, as described in Sambrook et al., supra, or electroporation is generally used for prokaryotes. Infection with *Agrobacterium tumefaciens* is used for transformation of certain plant cells, as described by Shaw et al., Gene, 23:315 (1983) and WO 89/05859 published 29 June 1989. For mammalian cells without such cell walls, the calcium phosphate precipitation method of Graham and van der Eb, Virology, 52:456-457 (1978) can be employed. General aspects of mammalian cell host system transfections have been described in U.S. Patent No. 4,399,216. Transformations into yeast are typically carried out according to the method of Van Solingen et al., J. Bact., 130:946 (1977) and Hsiao et al., Proc. Natl. Acad. Sci. (USA), 76:3829 (1979). However, other methods for introducing DNA into cells, such as by nuclear microinjection, electroporation, bacterial protoplast fusion with intact cells, or polycations, e.g., polybrene, polyornithine, may also be used. For various techniques for transforming mammalian cells, see Keown et al., Methods in Enzymology, 185:527-537 (1990) and Mansour et al., Nature, 336:348-352 (1988).

Suitable host cells for cloning or expressing the DNA in the vectors herein include prokaryote, yeast, or higher eukaryote cells. Suitable prokaryotes include but are not limited to eubacteria, such as Gram-negative or Gram-positive organisms, for example, Enterobacteriaceae such as *E. coli*. Various *E. coli* strains are publicly available, such as *E. coli* K12 strain MM294 (ATCC 31,446); *E. coli* X1776 (ATCC 31,537); *E. coli* strain W3110 (ATCC 27,325) and K5 772 (ATCC 53,635). Other suitable prokaryotic host cells include Enterobacteriaceae such as *Escherichia*, e.g., *E. coli*, *Enterobacter*, *Erwinia*, *Klebsiella*, *Proteus*, *Salmonella*, e.g., *Salmonella typhimurium*, *Serratia*, e.g., *Serratia marcescans*, and *Shigella*, as well as *Bacilli* such as *B. subtilis* and *B. licheniformis* (e.g., *B. licheniformis* 41P disclosed in DD 266,710 published 12 April 1989), *Pseudomonas* such as *P. aeruginosa*, and *Streptomyces*. These examples are illustrative rather than limiting. Strain W3110 is one particularly preferred host or parent host because it is a common host strain for recombinant DNA product fermentations. Preferably, the host cell secretes minimal amounts of proteolytic enzymes. For example, strain W3110 may be modified to effect a genetic mutation in the genes encoding proteins endogenous to the host, with examples of such hosts including *E. coli* W3110 strain 1A2, which has the complete genotype *tonA*; *E. coli* W3110 strain 9E4, which has the complete genotype *tonA ptr3*; *E. coli* W3110 strain 27C7 (ATCC



55,244), which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT kan<sup>r</sup>*; *E. coli* W3110 strain 37D6, which has the complete genotype *tonA ptr3 phoA E15 (argF-lac)169 degP ompT rbs7 ilvG kan<sup>r</sup>*; *E. coli* W3110 strain 40B4, which is strain 37D6 with a non-kanamycin resistant *degP* deletion mutation; and an *E. coli* strain having mutant periplasmic protease disclosed in U.S. Patent No. 4,946,783 issued 7 August 1990. Alternatively, *in vitro* methods of cloning, e.g., PCR or other nucleic acid polymerase reactions, are suitable.

5 In addition to prokaryotes, eukaryotic microbes such as filamentous fungi or yeast are suitable cloning or expression hosts for PRO-encoding vectors. *Saccharomyces cerevisiae* is a commonly used lower eukaryotic host microorganism. Others include *Schizosaccharomyces pombe* (Beach and Nurse, Nature, 290: 140 [1981]; EP 139,383 published 2 May 1985); *Kluyveromyces* hosts (U.S. Patent No. 4,943,529; Fleer et al., Bio/Technology, 9:968-975 (1991)) such as, e.g., *K. lactis* (MW98-8C, CBS683, CBS4574; Louvencourt et al.,  
10 J. Bacteriol., 154(2):737-742 [1983]), *K. fragilis* (ATCC 12,424), *K. bulgaricus* (ATCC 16,045), *K. wickerhamii* (ATCC 24,178), *K. waltii* (ATCC 56,500), *K. drosophilae* (ATCC 36,906; Van den Berg et al., Bio/Technology, 8:135 (1990)), *K. thermotolerans*, and *K. marxianus*; *yarrowia* (EP 402,226); *Pichia pastoris* (EP 183,070; Sreekrishna et al., J. Basic Microbiol., 28:265-278 [1988]); *Candida*; *Trichoderma reesia* (EP 244,234); *Neurospora crassa* (Case et al., Proc. Natl. Acad. Sci. USA, 76:5259-5263 [1979]); *Schwanniomyces*  
15 such as *Schwanniomyces occidentalis* (EP 394,538 published 31 October 1990); and filamentous fungi such as, e.g., *Neurospora*, *Penicillium*, *Tolypocladium* (WO 91/00357 published 10 January 1991), and *Aspergillus* hosts such as *A. nidulans* (Ballance et al., Biochem. Biophys. Res. Commun., 112:284-289 [1983]; Tilburn et al., Gene, 26:205-221 [1983]; Yelton et al., Proc. Natl. Acad. Sci. USA, 81: 1470-1474 [1984]) and *A. niger* (Kelly and Hynes, EMBO J., 4:475-479 [1985]). Methylophilic yeasts are suitable herein and include, but are not  
20 limited to, yeast capable of growth on methanol selected from the genera consisting of *Hansenula*, *Candida*, *Kloeckera*, *Pichia*, *Saccharomyces*, *Torulopsis*, and *Rhodotorula*. A list of specific species that are exemplary of this class of yeasts may be found in C. Anthony, The Biochemistry of Methylophilic Yeasts, 269 (1982).

Suitable host cells for the expression of glycosylated PRO are derived from multicellular organisms. Examples of invertebrate cells include insect cells such as *Drosophila* S2 and *Spodoptera* Sf9, as well as plant  
25 cells. Examples of useful mammalian host cell lines include Chinese hamster ovary (CHO) and COS cells. More specific examples include monkey kidney CV1 line transformed by SV40 (COS-7, ATCC CRL 1651); human embryonic kidney line (293 or 293 cells subcloned for growth in suspension culture, Graham et al., J. Gen. Virol., 36:59 (1977)); Chinese hamster ovary cells/-DHFR (CHO, Urlaub and Chasin, Proc. Natl. Acad. Sci. USA, 77:4216 (1980)); mouse sertoli cells (TM4, Mather, Biol. Reprod., 23:243-251 (1980)); human lung cells (W138,  
30 ATCC CCL 75); human liver cells (Hep G2, HB 8065); and mouse mammary tumor (MMT 060562, ATCC CCL51). The selection of the appropriate host cell is deemed to be within the skill in the art.

### 3. Selection and Use of a Replicable Vector

35 The nucleic acid (e.g., cDNA or genomic DNA) encoding PRO may be inserted into a replicable vector for cloning (amplification of the DNA) or for expression. Various vectors are publicly available. The vector may, for example, be in the form of a plasmid, cosmid, viral particle, or phage. The appropriate nucleic acid sequence may be inserted into the vector by a variety of procedures. In general, DNA is inserted into an

appropriate restriction endonuclease site(s) using techniques known in the art. Vector components generally include, but are not limited to, one or more of a signal sequence, an origin of replication, one or more marker genes, an enhancer element, a promoter, and a transcription termination sequence. Construction of suitable vectors containing one or more of these components employs standard ligation techniques which are known to the skilled artisan.

5           The PRO may be produced recombinantly not only directly, but also as a fusion polypeptide with a heterologous polypeptide, which may be a signal sequence or other polypeptide having a specific cleavage site at the N-terminus of the mature protein or polypeptide. In general, the signal sequence may be a component of the vector, or it may be a part of the PRO-encoding DNA that is inserted into the vector. The signal sequence may be a prokaryotic signal sequence selected, for example, from the group of the alkaline phosphatase, penicillinase, lpp, or heat-stable enterotoxin II leaders. For yeast secretion the signal sequence may be, e.g., the yeast invertase leader, alpha factor leader (including *Saccharomyces* and *Kluyveromyces*  $\alpha$ -factor leaders, the latter described in U.S. Patent No. 5,010,182), or acid phosphatase leader, the *C. albicans* glucoamylase leader (EP 362,179 published 4 April 1990), or the signal described in WO 90/13646 published 15 November 1990. In mammalian cell expression, mammalian signal sequences may be used to direct secretion of the protein, such as  
10           signal sequences from secreted polypeptides of the same or related species, as well as viral secretory leaders.

15           Both expression and cloning vectors contain a nucleic acid sequence that enables the vector to replicate in one or more selected host cells. Such sequences are well known for a variety of bacteria, yeast, and viruses. The origin of replication from the plasmid pBR322 is suitable for most Gram-negative bacteria, the 2 $\mu$  plasmid origin is suitable for yeast, and various viral origins (SV40, polyoma, adenovirus, VSV or BPV) are useful for  
20           cloning vectors in mammalian cells.

            Expression and cloning vectors will typically contain a selection gene, also termed a selectable marker. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxins, e.g., ampicillin, neomycin, methotrexate, or tetracycline, (b) complement auxotrophic deficiencies, or (c) supply critical nutrients not available from complex media, e.g., the gene encoding D-alanine racemase for *Bacilli*.

25           An example of suitable selectable markers for mammalian cells are those that enable the identification of cells competent to take up the PRO-encoding nucleic acid, such as DHFR or thymidine kinase. An appropriate host cell when wild-type DHFR is employed is the CHO cell line deficient in DHFR activity, prepared and propagated as described by Urlaub et al., Proc. Natl. Acad. Sci. USA, 77:4216 (1980). A suitable selection gene for use in yeast is the *trp1* gene present in the yeast plasmid YRp7 [Stinchcomb et al., Nature, 282:39 (1979); Kingsman et al., Gene, 7:141 (1979); Tschemper et al., Gene, 10:157 (1980)]. The *trp1* gene provides a  
30           selection marker for a mutant strain of yeast lacking the ability to grow in tryptophan, for example, ATCC No. 44076 or PEP4-1 [Jones, Genetics, 85:12 (1977)].

            Expression and cloning vectors usually contain a promoter operably linked to the PRO-encoding nucleic acid sequence to direct mRNA synthesis. Promoters recognized by a variety of potential host cells are well  
35           known. Promoters suitable for use with prokaryotic hosts include the  $\beta$ -lactamase and lactose promoter systems [Chang et al., Nature, 275:615 (1978); Goeddel et al., Nature, 281:544 (1979)], alkaline phosphatase, a tryptophan (*trp*) promoter system [Goeddel, Nucleic Acids Res., 8:4057 (1980); EP 36,776], and hybrid

promoters such as the tac promoter [deBoer et al., Proc. Natl. Acad. Sci. USA, 80:21-25 (1983)]. Promoters for use in bacterial systems also will contain a Shine-Dalgarno (S.D.) sequence operably linked to the DNA encoding PRO.

Examples of suitable promoting sequences for use with yeast hosts include the promoters for 3-phosphoglycerate kinase [Hitzeman et al., J. Biol. Chem., 255:2073 (1980)] or other glycolytic enzymes [Hess et al., J. Adv. Enzyme Reg., 7:149 (1968); Holland, Biochemistry, 17:4900 (1978)], such as enolase, glyceraldehyde-3-phosphate dehydrogenase, hexokinase, pyruvate decarboxylase, phosphofructokinase, glucose-6-phosphate isomerase, 3-phosphoglycerate mutase, pyruvate kinase, triosephosphate isomerase, phosphoglucose isomerase, and glucokinase.

Other yeast promoters, which are inducible promoters having the additional advantage of transcription controlled by growth conditions, are the promoter regions for alcohol dehydrogenase 2, isocytochrome C, acid phosphatase, degradative enzymes associated with nitrogen metabolism, metallothionein, glyceraldehyde-3-phosphate dehydrogenase, and enzymes responsible for maltose and galactose utilization. Suitable vectors and promoters for use in yeast expression are further described in EP 73,657.

PRO transcription from vectors in mammalian host cells is controlled, for example, by promoters obtained from the genomes of viruses such as polyoma virus, fowlpox virus (UK 2,211,504 published 5 July 1989), adenovirus (such as Adenovirus 2), bovine papilloma virus, avian sarcoma virus, cytomegalovirus, a retrovirus, hepatitis-B virus and Simian Virus 40 (SV40), from heterologous mammalian promoters, e.g., the actin promoter or an immunoglobulin promoter, and from heat-shock promoters, provided such promoters are compatible with the host cell systems.

Transcription of a DNA encoding the PRO by higher eukaryotes may be increased by inserting an enhancer sequence into the vector. Enhancers are cis-acting elements of DNA, usually about from 10 to 300 bp, that act on a promoter to increase its transcription. Many enhancer sequences are now known from mammalian genes (globin, elastase, albumin,  $\alpha$ -fetoprotein, and insulin). Typically, however, one will use an enhancer from a eukaryotic cell virus. Examples include the SV40 enhancer on the late side of the replication origin (bp 100-270), the cytomegalovirus early promoter enhancer, the polyoma enhancer on the late side of the replication origin, and adenovirus enhancers. The enhancer may be spliced into the vector at a position 5' or 3' to the PRO coding sequence, but is preferably located at a site 5' from the promoter.

Expression vectors used in eukaryotic host cells (yeast, fungi, insect, plant, animal, human, or nucleated cells from other multicellular organisms) will also contain sequences necessary for the termination of transcription and for stabilizing the mRNA. Such sequences are commonly available from the 5' and, occasionally 3', untranslated regions of eukaryotic or viral DNAs or cDNAs. These regions contain nucleotide segments transcribed as polyadenylated fragments in the untranslated portion of the mRNA encoding PRO.

Still other methods, vectors, and host cells suitable for adaptation to the synthesis of PRO in recombinant vertebrate cell culture are described in Gething et al., Nature, 293:620-625 (1981); Mantei et al., Nature, 281:40-46 (1979); EP 117,060; and EP 117,058.

#### 4. Detecting Gene Amplification/Expression

Gene amplification and/or expression may be measured in a sample directly, for example, by conventional Southern blotting, Northern blotting to quantitate the transcription of mRNA [Thomas, Proc. Natl. Acad. Sci. USA, 77:5201-5205 (1980)], dot blotting (DNA analysis), or *in situ* hybridization, using an appropriately labeled probe, based on the sequences provided herein. Alternatively, antibodies may be employed that can recognize specific duplexes, including DNA duplexes, RNA duplexes, and DNA-RNA hybrid duplexes or DNA-protein duplexes. The antibodies in turn may be labeled and the assay may be carried out where the duplex is bound to a surface, so that upon the formation of duplex on the surface, the presence of antibody bound to the duplex can be detected.

Gene expression, alternatively, may be measured by immunological methods, such as immunohistochemical staining of cells or tissue sections and assay of cell culture or body fluids, to quantitate directly the expression of gene product. Antibodies useful for immunohistochemical staining and/or assay of sample fluids may be either monoclonal or polyclonal, and may be prepared in any mammal. Conveniently, the antibodies may be prepared against a native sequence PRO polypeptide or against a synthetic peptide based on the DNA sequences provided herein or against exogenous sequence fused to PRO DNA and encoding a specific antibody epitope.

#### 5. Purification of Polypeptide

Forms of PRO may be recovered from culture medium or from host cell lysates. If membrane-bound, it can be released from the membrane using a suitable detergent solution (e.g. Triton-X 100) or by enzymatic cleavage. Cells employed in expression of PRO can be disrupted by various physical or chemical means, such as freeze-thaw cycling, sonication, mechanical disruption, or cell lysing agents.

It may be desired to purify PRO from recombinant cell proteins or polypeptides. The following procedures are exemplary of suitable purification procedures: by fractionation on an ion-exchange column; ethanol precipitation; reverse phase HPLC; chromatography on silica or on a cation-exchange resin such as DEAE; chromatofocusing; SDS-PAGE; ammonium sulfate precipitation; gel filtration using, for example, Sephadex G-75; protein A Sepharose columns to remove contaminants such as IgG; and metal chelating columns to bind epitope-tagged forms of the PRO. Various methods of protein purification may be employed and such methods are known in the art and described for example in Deutscher, Methods in Enzymology, 182 (1990); Scopes, Protein Purification: Principles and Practice, Springer-Verlag, New York (1982). The purification step(s) selected will depend, for example, on the nature of the production process used and the particular PRO produced.

#### E. Uses for PRO

Nucleotide sequences (or their complement) encoding PRO have various applications in the art of molecular biology, including uses as hybridization probes, in chromosome and gene mapping and in the generation of anti-sense RNA and DNA. PRO nucleic acid will also be useful for the preparation of PRO polypeptides by the recombinant techniques described herein.

The full-length native sequence PRO gene, or portions thereof, may be used as hybridization probes for a cDNA library to isolate the full-length PRO cDNA or to isolate still other cDNAs (for instance, those encoding

naturally-occurring variants of PRO or PRO from other species) which have a desired sequence identity to the native PRO sequence disclosed herein. Optionally, the length of the probes will be about 20 to about 50 bases. The hybridization probes may be derived from at least partially novel regions of the full length native nucleotide sequence wherein those regions may be determined without undue experimentation or from genomic sequences including promoters, enhancer elements and introns of native sequence PRO. By way of example, a screening method will comprise isolating the coding region of the PRO gene using the known DNA sequence to synthesize a selected probe of about 40 bases. Hybridization probes may be labeled by a variety of labels, including radionucleotides such as  $^{32}\text{P}$  or  $^{35}\text{S}$ , or enzymatic labels such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems. Labeled probes having a sequence complementary to that of the PRO gene of the present invention can be used to screen libraries of human cDNA, genomic DNA or mRNA to determine which members of such libraries the probe hybridizes to. Hybridization techniques are described in further detail in the Examples below.

Any EST sequences disclosed in the present application may similarly be employed as probes, using the methods disclosed herein.

Other useful fragments of the PRO nucleic acids include antisense or sense oligonucleotides comprising a single-stranded nucleic acid sequence (either RNA or DNA) capable of binding to target PRO mRNA (sense) or PRO DNA (antisense) sequences. Antisense or sense oligonucleotides, according to the present invention, comprise a fragment of the coding region of PRO DNA. Such a fragment generally comprises at least about 14 nucleotides, preferably from about 14 to 30 nucleotides. The ability to derive an antisense or a sense oligonucleotide, based upon a cDNA sequence encoding a given protein is described in, for example, Stein and Cohen (Cancer Res. 48:2659, 1988) and van der Krol et al. (BioTechniques 6:958, 1988).

Binding of antisense or sense oligonucleotides to target nucleic acid sequences results in the formation of duplexes that block transcription or translation of the target sequence by one of several means, including enhanced degradation of the duplexes, premature termination of transcription or translation, or by other means. The antisense oligonucleotides thus may be used to block expression of PRO proteins. Antisense or sense oligonucleotides further comprise oligonucleotides having modified sugar-phosphodiester backbones (or other sugar linkages, such as those described in WO 91/06629) and wherein such sugar linkages are resistant to endogenous nucleases. Such oligonucleotides with resistant sugar linkages are stable *in vivo* (i.e., capable of resisting enzymatic degradation) but retain sequence specificity to be able to bind to target nucleotide sequences.

Other examples of sense or antisense oligonucleotides include those oligonucleotides which are covalently linked to organic moieties, such as those described in WO 90/10048, and other moieties that increases affinity of the oligonucleotide for a target nucleic acid sequence, such as poly-(L-lysine). Further still, intercalating agents, such as ellipticine, and alkylating agents or metal complexes may be attached to sense or antisense oligonucleotides to modify binding specificities of the antisense or sense oligonucleotide for the target nucleotide sequence.

Antisense or sense oligonucleotides may be introduced into a cell containing the target nucleic acid sequence by any gene transfer method, including, for example,  $\text{CaPO}_4$ -mediated DNA transfection, electroporation, or by using gene transfer vectors such as Epstein-Barr virus. In a preferred procedure, an

antisense or sense oligonucleotide is inserted into a suitable retroviral vector. A cell containing the target nucleic acid sequence is contacted with the recombinant retroviral vector, either *in vivo* or *ex vivo*. Suitable retroviral vectors include, but are not limited to, those derived from the murine retrovirus M-MuLV, N2 (a retrovirus derived from M-MuLV), or the double copy vectors designated DCT5A, DCT5B and DCT5C (see WO 90/13641).

5           Sense or antisense oligonucleotides also may be introduced into a cell containing the target nucleotide sequence by formation of a conjugate with a ligand binding molecule, as described in WO 91/04753. Suitable ligand binding molecules include, but are not limited to, cell surface receptors, growth factors, other cytokines, or other ligands that bind to cell surface receptors. Preferably, conjugation of the ligand binding molecule does not substantially interfere with the ability of the ligand binding molecule to bind to its corresponding molecule or  
10 receptor, or block entry of the sense or antisense oligonucleotide or its conjugated version into the cell.

Alternatively, a sense or an antisense oligonucleotide may be introduced into a cell containing the target nucleic acid sequence by formation of an oligonucleotide-lipid complex, as described in WO 90/10448. The sense or antisense oligonucleotide-lipid complex is preferably dissociated within the cell by an endogenous lipase.

15           Antisense or sense RNA or DNA molecules are generally at least about 5 bases in length, about 10 bases in length, about 15 bases in length, about 20 bases in length, about 25 bases in length, about 30 bases in length, about 35 bases in length, about 40 bases in length, about 45 bases in length, about 50 bases in length, about 55 bases in length, about 60 bases in length, about 65 bases in length, about 70 bases in length, about 75 bases in length, about 80 bases in length, about 85 bases in length, about 90 bases in length, about 95 bases in length, about 100 bases in length, or more.

20           The probes may also be employed in PCR techniques to generate a pool of sequences for identification of closely related PRO coding sequences.

Nucleotide sequences encoding a PRO can also be used to construct hybridization probes for mapping the gene which encodes that PRO and for the genetic analysis of individuals with genetic disorders. The nucleotide sequences provided herein may be mapped to a chromosome and specific regions of a chromosome  
25 using known techniques, such as *in situ* hybridization, linkage analysis against known chromosomal markers, and hybridization screening with libraries.

When the coding sequences for PRO encode a protein which binds to another protein (example, where the PRO is a receptor), the PRO can be used in assays to identify the other proteins or molecules involved in the binding interaction. By such methods, inhibitors of the receptor/ligand binding interaction can be identified.  
30 Proteins involved in such binding interactions can also be used to screen for peptide or small molecule inhibitors or agonists of the binding interaction. Also, the receptor PRO can be used to isolate correlative ligand(s). Screening assays can be designed to find lead compounds that mimic the biological activity of a native PRO or a receptor for PRO. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates. Small molecules  
35 contemplated include synthetic organic or inorganic compounds. The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays and cell based assays, which are well characterized in the art.

Nucleic acids which encode PRO or its modified forms can also be used to generate either transgenic animals or "knock out" animals which, in turn, are useful in the development and screening of therapeutically useful reagents. A transgenic animal (e.g., a mouse or rat) is an animal having cells that contain a transgene, which transgene was introduced into the animal or an ancestor of the animal at a prenatal, e.g., an embryonic stage. A transgene is a DNA which is integrated into the genome of a cell from which a transgenic animal develops. In one embodiment, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques and the genomic sequences used to generate transgenic animals that contain cells which express DNA encoding PRO. Methods for generating transgenic animals, particularly animals such as mice or rats, have become conventional in the art and are described, for example, in U.S. Patent Nos. 4,736,866 and 4,870,009. Typically, particular cells would be targeted for PRO transgene incorporation with tissue-specific enhancers. Transgenic animals that include a copy of a transgene encoding PRO introduced into the germ line of the animal at an embryonic stage can be used to examine the effect of increased expression of DNA encoding PRO. Such animals can be used as tester animals for reagents thought to confer protection from, for example, pathological conditions associated with its overexpression. In accordance with this facet of the invention, an animal is treated with the reagent and a reduced incidence of the pathological condition, compared to untreated animals bearing the transgene, would indicate a potential therapeutic intervention for the pathological condition.

Alternatively, non-human homologues of PRO can be used to construct a PRO "knock out" animal which has a defective or altered gene encoding PRO as a result of homologous recombination between the endogenous gene encoding PRO and altered genomic DNA encoding PRO introduced into an embryonic stem cell of the animal. For example, cDNA encoding PRO can be used to clone genomic DNA encoding PRO in accordance with established techniques. A portion of the genomic DNA encoding PRO can be deleted or replaced with another gene, such as a gene encoding a selectable marker which can be used to monitor integration. Typically, several kilobases of unaltered flanking DNA (both at the 5' and 3' ends) are included in the vector [see e.g., Thomas and Capecchi, Cell, 51:503 (1987) for a description of homologous recombination vectors]. The vector is introduced into an embryonic stem cell line (e.g., by electroporation) and cells in which the introduced DNA has homologously recombined with the endogenous DNA are selected [see e.g., Li et al., Cell, 69:915 (1992)]. The selected cells are then injected into a blastocyst of an animal (e.g., a mouse or rat) to form aggregation chimeras [see e.g., Bradley, in *Teratocarcinomas and Embryonic Stem Cells: A Practical Approach*, E. J. Robertson, ed. (IRL, Oxford, 1987), pp. 113-152]. A chimeric embryo can then be implanted into a suitable pseudopregnant female foster animal and the embryo brought to term to create a "knock out" animal. Progeny harboring the homologously recombined DNA in their germ cells can be identified by standard techniques and used to breed animals in which all cells of the animal contain the homologously recombined DNA. Knockout animals can be characterized for instance, for their ability to defend against certain pathological conditions and for their development of pathological conditions due to absence of the PRO polypeptide.

Nucleic acid encoding the PRO polypeptides may also be used in gene therapy. In gene therapy applications, genes are introduced into cells in order to achieve *in vivo* synthesis of a therapeutically effective genetic product, for example for replacement of a defective gene. "Gene therapy" includes both conventional

gene therapy where a lasting effect is achieved by a single treatment, and the administration of gene therapeutic agents, which involves the one time or repeated administration of a therapeutically effective DNA or mRNA. Antisense RNAs and DNAs can be used as therapeutic agents for blocking the expression of certain genes *in vivo*. It has already been shown that short antisense oligonucleotides can be imported into cells where they act as inhibitors, despite their low intracellular concentrations caused by their restricted uptake by the cell membrane.

5 (Zamecnik *et al.*, Proc. Natl. Acad. Sci. USA 83:4143-4146 [1986]). The oligonucleotides can be modified to enhance their uptake, e.g. by substituting their negatively charged phosphodiester groups by uncharged groups.

There are a variety of techniques available for introducing nucleic acids into viable cells. The techniques vary depending upon whether the nucleic acid is transferred into cultured cells *in vitro*, or *in vivo* in the cells of the intended host. Techniques suitable for the transfer of nucleic acid into mammalian cells *in vitro* include the use of liposomes, electroporation, microinjection, cell fusion, DEAE-dextran, the calcium phosphate precipitation

10 method, etc. The currently preferred *in vivo* gene transfer techniques include transfection with viral (typically retroviral) vectors and viral coat protein-liposome mediated transfection (Dzau *et al.*, Trends in Biotechnology 11, 205-210 [1993]). In some situations it is desirable to provide the nucleic acid source with an agent that targets the target cells, such as an antibody specific for a cell surface membrane protein or the target cell, a ligand for a receptor on the target cell, etc. Where liposomes are employed, proteins which bind to a cell surface membrane protein associated with endocytosis may be used for targeting and/or to facilitate uptake, e.g. capsid proteins or fragments thereof tropic for a particular cell type, antibodies for proteins which undergo internalization in cycling, proteins that target intracellular localization and enhance intracellular half-life. The technique of receptor-mediated endocytosis is described, for example, by Wu *et al.*, J. Biol. Chem. 262, 4429-4432 (1987); and Wagner

15 *et al.*, Proc. Natl. Acad. Sci. USA 87, 3410-3414 (1990). For review of gene marking and gene therapy protocols see Anderson *et al.*, Science 256, 808-813 (1992).

The PRO polypeptides described herein may also be employed as molecular weight markers for protein electrophoresis purposes and the isolated nucleic acid sequences may be used for recombinantly expressing those markers.

25 The nucleic acid molecules encoding the PRO polypeptides or fragments thereof described herein are useful for chromosome identification. In this regard, there exists an ongoing need to identify new chromosome markers, since relatively few chromosome marking reagents, based upon actual sequence data are presently available. Each PRO nucleic acid molecule of the present invention can be used as a chromosome marker.

The PRO polypeptides and nucleic acid molecules of the present invention may also be used

30 diagnostically for tissue typing, wherein the PRO polypeptides of the present invention may be differentially expressed in one tissue as compared to another, preferably in a diseased tissue as compared to a normal tissue of the same tissue type. PRO nucleic acid molecules will find use for generating probes for PCR, Northern analysis, Southern analysis and Western analysis.

The PRO polypeptides described herein may also be employed as therapeutic agents. The PRO

35 polypeptides of the present invention can be formulated according to known methods to prepare pharmaceutically useful compositions, whereby the PRO product hereof is combined in admixture with a pharmaceutically acceptable carrier vehicle. Therapeutic formulations are prepared for storage by mixing the active ingredient



having the desired degree of purity with optional physiologically acceptable carriers, excipients or stabilizers (Remington's Pharmaceutical Sciences 16th edition, Osol, A. Ed. (1980)), in the form of lyophilized formulations or aqueous solutions. Acceptable carriers, excipients or stabilizers are nontoxic to recipients at the dosages and concentrations employed, and include buffers such as phosphate, citrate and other organic acids; antioxidants including ascorbic acid; low molecular weight (less than about 10 residues) polypeptides; proteins, such as serum  
5 albumin, gelatin or immunoglobulins; hydrophilic polymers such as polyvinylpyrrolidone, amino acids such as glycine, glutamine, asparagine, arginine or lysine; monosaccharides, disaccharides and other carbohydrates including glucose, mannose, or dextrans; chelating agents such as EDTA; sugar alcohols such as mannitol or sorbitol; salt-forming counterions such as sodium; and/or nonionic surfactants such as TWEEN<sup>TM</sup>, PLURONICS<sup>TM</sup> or PEG.

10 The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes, prior to or following lyophilization and reconstitution.

Therapeutic compositions herein generally are placed into a container having a sterile access port, for example, an intravenous solution bag or vial having a stopper pierceable by a hypodermic injection needle.

The route of administration is in accord with known methods, e.g. injection or infusion by intravenous,  
15 intraperitoneal, intracerebral, intramuscular, intraocular, intraarterial or intralesional routes, topical administration, or by sustained release systems.

Dosages and desired drug concentrations of pharmaceutical compositions of the present invention may vary depending on the particular use envisioned. The determination of the appropriate dosage or route of administration is well within the skill of an ordinary physician. Animal experiments provide reliable guidance  
20 for the determination of effective doses for human therapy. Interspecies scaling of effective doses can be performed following the principles laid down by Mordenti, J. and Chappell, W. "The use of interspecies scaling in toxicokinetics" In *Toxicokinetics and New Drug Development*, Yacobi et al., Eds., Pergamon Press, New York 1989, pp. 42-96.

When *in vivo* administration of a PRO polypeptide or agonist or antagonist thereof is employed, normal  
25 dosage amounts may vary from about 10 ng/kg to up to 100 mg/kg of mammal body weight or more per day, preferably about 1 µg/kg/day to 10 mg/kg/day, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature; see, for example, U.S. Pat. Nos. 4,657,760; 5,206,344; or 5,225,212. It is anticipated that different formulations will be effective for different treatment compounds and different disorders, that administration targeting one organ or tissue, for example, may  
30 necessitate delivery in a manner different from that to another organ or tissue.

Where sustained-release administration of a PRO polypeptide is desired in a formulation with release characteristics suitable for the treatment of any disease or disorder requiring administration of the PRO polypeptide, microencapsulation of the PRO polypeptide is contemplated. Microencapsulation of recombinant proteins for sustained release has been successfully performed with human growth hormone (rhGH), interferon-  
35 (rhIFN-), interleukin-2, and MN rgp120. Johnson et al., Nat. Med., 2:795-799 (1996); Yasuda, Biomed. Ther., 27:1221-1223 (1993); Hora et al., Bio/Technology, 8:755-758 (1990); Cleland, "Design and Production of Single Immunization Vaccines Using Polylactide Polyglycolide Microsphere Systems," in Vaccine Design: The Subunit

and Adjuvant Approach, Powell and Newman, eds, (Plenum Press: New York, 1995), pp. 439-462; WO 97/03692, WO 96/40072, WO 96/07399; and U.S. Pat. No. 5,654,010.

The sustained-release formulations of these proteins were developed using poly-lactic-coglycolic acid (PLGA) polymer due to its biocompatibility and wide range of biodegradable properties. The degradation products of PLGA, lactic and glycolic acids, can be cleared quickly within the human body. Moreover, the degradability of this polymer can be adjusted from months to years depending on its molecular weight and composition. Lewis, "Controlled release of bioactive agents from lactide/glycolide polymer," in: M. Chasin and R. Langer (Eds.), Biodegradable Polymers as Drug Delivery Systems (Marcel Dekker: New York, 1990), pp. 1-41.

This invention encompasses methods of screening compounds to identify those that mimic the PRO polypeptide (agonists) or prevent the effect of the PRO polypeptide (antagonists). Screening assays for antagonist drug candidates are designed to identify compounds that bind or complex with the PRO polypeptides encoded by the genes identified herein, or otherwise interfere with the interaction of the encoded polypeptides with other cellular proteins. Such screening assays will include assays amenable to high-throughput screening of chemical libraries, making them particularly suitable for identifying small molecule drug candidates.

The assays can be performed in a variety of formats, including protein-protein binding assays, biochemical screening assays, immunoassays, and cell-based assays, which are well characterized in the art.

All assays for antagonists are common in that they call for contacting the drug candidate with a PRO polypeptide encoded by a nucleic acid identified herein under conditions and for a time sufficient to allow these two components to interact.

In binding assays, the interaction is binding and the complex formed can be isolated or detected in the reaction mixture. In a particular embodiment, the PRO polypeptide encoded by the gene identified herein or the drug candidate is immobilized on a solid phase, e.g., on a microtiter plate, by covalent or non-covalent attachments. Non-covalent attachment generally is accomplished by coating the solid surface with a solution of the PRO polypeptide and drying. Alternatively, an immobilized antibody, e.g., a monoclonal antibody, specific for the PRO polypeptide to be immobilized can be used to anchor it to a solid surface. The assay is performed by adding the non-immobilized component, which may be labeled by a detectable label, to the immobilized component, e.g., the coated surface containing the anchored component. When the reaction is complete, the non-reacted components are removed, e.g., by washing, and complexes anchored on the solid surface are detected. When the originally non-immobilized component carries a detectable label, the detection of label immobilized on the surface indicates that complexing occurred. Where the originally non-immobilized component does not carry a label, complexing can be detected, for example, by using a labeled antibody specifically binding the immobilized complex.

If the candidate compound interacts with but does not bind to a particular PRO polypeptide encoded by a gene identified herein, its interaction with that polypeptide can be assayed by methods well known for detecting protein-protein interactions. Such assays include traditional approaches, such as, e.g., cross-linking, co-immunoprecipitation, and co-purification through gradients or chromatographic columns. In addition, protein-protein interactions can be monitored by using a yeast-based genetic system described by Fields and co-workers (Fields and Song, Nature (London), 340:245-246 (1989); Chien et al., Proc. Natl. Acad. Sci. USA, 88:9578-9582

(1991)) as disclosed by Chevray and Nathans, Proc. Natl. Acad. Sci. USA, 89: 5789-5793 (1991). Many transcriptional activators, such as yeast GAL4, consist of two physically discrete modular domains, one acting as the DNA-binding domain, the other one functioning as the transcription-activation domain. The yeast expression system described in the foregoing publications (generally referred to as the "two-hybrid system") takes advantage of this property, and employs two hybrid proteins, one in which the target protein is fused to the DNA-binding domain of GAL4, and another, in which candidate activating proteins are fused to the activation domain. The expression of a GAL1-*lacZ* reporter gene under control of a GAL4-activated promoter depends on reconstitution of GAL4 activity via protein-protein interaction. Colonies containing interacting polypeptides are detected with a chromogenic substrate for  $\beta$ -galactosidase. A complete kit (MATCHMAKER™) for identifying protein-protein interactions between two specific proteins using the two-hybrid technique is commercially available from Clontech. This system can also be extended to map protein domains involved in specific protein interactions as well as to pinpoint amino acid residues that are crucial for these interactions.

Compounds that interfere with the interaction of a gene encoding a PRO polypeptide identified herein and other intra- or extracellular components can be tested as follows: usually a reaction mixture is prepared containing the product of the gene and the intra- or extracellular component under conditions and for a time allowing for the interaction and binding of the two products. To test the ability of a candidate compound to inhibit binding, the reaction is run in the absence and in the presence of the test compound. In addition, a placebo may be added to a third reaction mixture, to serve as positive control. The binding (complex formation) between the test compound and the intra- or extracellular component present in the mixture is monitored as described hereinabove. The formation of a complex in the control reaction(s) but not in the reaction mixture containing the test compound indicates that the test compound interferes with the interaction of the test compound and its reaction partner.

To assay for antagonists, the PRO polypeptide may be added to a cell along with the compound to be screened for a particular activity and the ability of the compound to inhibit the activity of interest in the presence of the PRO polypeptide indicates that the compound is an antagonist to the PRO polypeptide. Alternatively, antagonists may be detected by combining the PRO polypeptide and a potential antagonist with membrane-bound PRO polypeptide receptors or recombinant receptors under appropriate conditions for a competitive inhibition assay. The PRO polypeptide can be labeled, such as by radioactivity, such that the number of PRO polypeptide molecules bound to the receptor can be used to determine the effectiveness of the potential antagonist. The gene encoding the receptor can be identified by numerous methods known to those of skill in the art, for example, ligand panning and FACS sorting. Coligan et al., Current Protocols in Immun., 1(2): Chapter 5 (1991). Preferably, expression cloning is employed wherein polyadenylated RNA is prepared from a cell responsive to the PRO polypeptide and a cDNA library created from this RNA is divided into pools and used to transfect COS cells or other cells that are not responsive to the PRO polypeptide. Transfected cells that are grown on glass slides are exposed to labeled PRO polypeptide. The PRO polypeptide can be labeled by a variety of means including iodination or inclusion of a recognition site for a site-specific protein kinase. Following fixation and incubation, the slides are subjected to autoradiographic analysis. Positive pools are identified and sub-pools are prepared and re-transfected using an interactive sub-pooling and re-screening process, eventually yielding a single

clone that encodes the putative receptor.

As an alternative approach for receptor identification, labeled PRO polypeptide can be photoaffinity-linked with cell membrane or extract preparations that express the receptor molecule. Cross-linked material is resolved by PAGE and exposed to X-ray film. The labeled complex containing the receptor can be excised, resolved into peptide fragments, and subjected to protein micro-sequencing. The amino acid sequence obtained from micro-sequencing would be used to design a set of degenerate oligonucleotide probes to screen a cDNA library to identify the gene encoding the putative receptor.

In another assay for antagonists, mammalian cells or a membrane preparation expressing the receptor would be incubated with labeled PRO polypeptide in the presence of the candidate compound. The ability of the compound to enhance or block this interaction could then be measured.

More specific examples of potential antagonists include an oligonucleotide that binds to the fusions of immunoglobulin with PRO polypeptide, and, in particular, antibodies including, without limitation, poly- and monoclonal antibodies and antibody fragments, single-chain antibodies, anti-idiotypic antibodies, and chimeric or humanized versions of such antibodies or fragments, as well as human antibodies and antibody fragments. Alternatively, a potential antagonist may be a closely related protein, for example, a mutated form of the PRO polypeptide that recognizes the receptor but imparts no effect, thereby competitively inhibiting the action of the PRO polypeptide.

Another potential PRO polypeptide antagonist is an antisense RNA or DNA construct prepared using antisense technology, where, e.g., an antisense RNA or DNA molecule acts to block directly the translation of mRNA by hybridizing to targeted mRNA and preventing protein translation. Antisense technology can be used to control gene expression through triple-helix formation or antisense DNA or RNA, both of which methods are based on binding of a polynucleotide to DNA or RNA. For example, the 5' coding portion of the polynucleotide sequence, which encodes the mature PRO polypeptides herein, is used to design an antisense RNA oligonucleotide of from about 10 to 40 base pairs in length. A DNA oligonucleotide is designed to be complementary to a region of the gene involved in transcription (triple helix - see Lee et al., Nucl. Acids Res., 6:3073 (1979); Cooney et al., Science, 241: 456 (1988); Dervan et al., Science, 251:1360 (1991)), thereby preventing transcription and the production of the PRO polypeptide. The antisense RNA oligonucleotide hybridizes to the mRNA *in vivo* and blocks translation of the mRNA molecule into the PRO polypeptide (antisense - Okano, Neurochem., 56:560 (1991); Oligodeoxynucleotides as Antisense Inhibitors of Gene Expression (CRC Press: Boca Raton, FL, 1988). The oligonucleotides described above can also be delivered to cells such that the antisense RNA or DNA may be expressed *in vivo* to inhibit production of the PRO polypeptide. When antisense DNA is used, oligodeoxyribonucleotides derived from the translation-initiation site, e.g., between about -10 and +10 positions of the target gene nucleotide sequence, are preferred.

Potential antagonists include small molecules that bind to the active site, the receptor binding site, or growth factor or other relevant binding site of the PRO polypeptide, thereby blocking the normal biological activity of the PRO polypeptide. Examples of small molecules include, but are not limited to, small peptides or peptide-like molecules, preferably soluble peptides, and synthetic non-peptidyl organic or inorganic compounds.

Ribozymes are enzymatic RNA molecules capable of catalyzing the specific cleavage of RNA.

Ribozymes act by sequence-specific hybridization to the complementary target RNA, followed by endonucleolytic cleavage. Specific ribozyme cleavage sites within a potential RNA target can be identified by known techniques. For further details see, e.g., Rossi, Current Biology, 4:469-471 (1994), and PCT publication No. WO 97/33551 (published September 18, 1997).

5 Nucleic acid molecules in triple-helix formation used to inhibit transcription should be single-stranded and composed of deoxynucleotides. The base composition of these oligonucleotides is designed such that it promotes triple-helix formation via Hoogsteen base-pairing rules, which generally require sizeable stretches of purines or pyrimidines on one strand of a duplex. For further details see, e.g., PCT publication No. WO 97/33551, *supra*.

10 These small molecules can be identified by any one or more of the screening assays discussed hereinabove and/or by any other screening techniques well known for those skilled in the art.

Diagnostic and therapeutic uses of the herein disclosed molecules may also be based upon the positive functional assay hits disclosed and described below.

#### F. Anti-PRO Antibodies

15 The present invention further provides anti-PRO antibodies. Exemplary antibodies include polyclonal, monoclonal, humanized, bispecific, and heteroconjugate antibodies.

##### 1. Polyclonal Antibodies

20 The anti-PRO antibodies may comprise polyclonal antibodies. Methods of preparing polyclonal antibodies are known to the skilled artisan. Polyclonal antibodies can be raised in a mammal, for example, by one or more injections of an immunizing agent and, if desired, an adjuvant. Typically, the immunizing agent and/or adjuvant will be injected in the mammal by multiple subcutaneous or intraperitoneal injections. The immunizing agent may include the PRO polypeptide or a fusion protein thereof. It may be useful to conjugate the immunizing agent to a protein known to be immunogenic in the mammal being immunized. Examples of such immunogenic proteins include but are not limited to keyhole limpet hemocyanin, serum albumin, bovine thyroglobulin, and soybean trypsin inhibitor. Examples of adjuvants which may be employed include Freund's complete adjuvant and MPL-TDM adjuvant (monophosphoryl Lipid A, synthetic trehalose dicorynomycolate).  
25 The immunization protocol may be selected by one skilled in the art without undue experimentation.

##### 30 2. Monoclonal Antibodies

The anti-PRO antibodies may, alternatively, be monoclonal antibodies. Monoclonal antibodies may be prepared using hybridoma methods, such as those described by Kohler and Milstein, Nature, 256:495 (1975). In a hybridoma method, a mouse, hamster, or other appropriate host animal, is typically immunized with an immunizing agent to elicit lymphocytes that produce or are capable of producing antibodies that will specifically bind to the immunizing agent. Alternatively, the lymphocytes may be immunized *in vitro*.  
35

The immunizing agent will typically include the PRO polypeptide or a fusion protein thereof. Generally, either peripheral blood lymphocytes ("PBLs") are used if cells of human origin are desired, or spleen cells or

lymph node cells are used if non-human mammalian sources are desired. The lymphocytes are then fused with an immortalized cell line using a suitable fusing agent, such as polyethylene glycol, to form a hybridoma cell [Goding, Monoclonal Antibodies: Principles and Practice, Academic Press, (1986) pp. 59-103]. Immortalized cell lines are usually transformed mammalian cells, particularly myeloma cells of rodent, bovine and human origin. Usually, rat or mouse myeloma cell lines are employed. The hybridoma cells may be cultured in a suitable culture medium that preferably contains one or more substances that inhibit the growth or survival of the unfused, immortalized cells. For example, if the parental cells lack the enzyme hypoxanthine guanine phosphoribosyl transferase (HGPRT or HPRT), the culture medium for the hybridomas typically will include hypoxanthine, aminopterin, and thymidine ("HAT medium"), which substances prevent the growth of HGPRT-deficient cells.

Preferred immortalized cell lines are those that fuse efficiently, support stable high level expression of antibody by the selected antibody-producing cells, and are sensitive to a medium such as HAT medium. More preferred immortalized cell lines are murine myeloma lines, which can be obtained, for instance, from the Salk Institute Cell Distribution Center, San Diego, California and the American Type Culture Collection, Manassas, Virginia. Human myeloma and mouse-human heteromyeloma cell lines also have been described for the production of human monoclonal antibodies [Kozbor, J. Immunol., 133:3001 (1984); Brodeur et al., Monoclonal Antibody Production Techniques and Applications, Marcel Dekker, Inc., New York, (1987) pp. 51-63].

The culture medium in which the hybridoma cells are cultured can then be assayed for the presence of monoclonal antibodies directed against PRO. Preferably, the binding specificity of monoclonal antibodies produced by the hybridoma cells is determined by immunoprecipitation or by an *in vitro* binding assay, such as radioimmunoassay (RIA) or enzyme-linked immunoabsorbent assay (ELISA). Such techniques and assays are known in the art. The binding affinity of the monoclonal antibody can, for example, be determined by the Scatchard analysis of Munson and Pollard, Anal. Biochem., 107:220 (1980).

After the desired hybridoma cells are identified, the clones may be subcloned by limiting dilution procedures and grown by standard methods [Goding, supra]. Suitable culture media for this purpose include, for example, Dulbecco's Modified Eagle's Medium and RPMI-1640 medium. Alternatively, the hybridoma cells may be grown *in vivo* as ascites in a mammal.

The monoclonal antibodies secreted by the subclones may be isolated or purified from the culture medium or ascites fluid by conventional immunoglobulin purification procedures such as, for example, protein A-Sepharose, hydroxylapatite chromatography, gel electrophoresis, dialysis, or affinity chromatography.

The monoclonal antibodies may also be made by recombinant DNA methods, such as those described in U.S. Patent No. 4,816,567. DNA encoding the monoclonal antibodies of the invention can be readily isolated and sequenced using conventional procedures (e.g., by using oligonucleotide probes that are capable of binding specifically to genes encoding the heavy and light chains of murine antibodies). The hybridoma cells of the invention serve as a preferred source of such DNA. Once isolated, the DNA may be placed into expression vectors, which are then transfected into host cells such as simian COS cells, Chinese hamster ovary (CHO) cells, or myeloma cells that do not otherwise produce immunoglobulin protein, to obtain the synthesis of monoclonal antibodies in the recombinant host cells. The DNA also may be modified, for example, by substituting the coding

sequence for human heavy and light chain constant domains in place of the homologous murine sequences [U.S. Patent No. 4,816,567; Morrison et al., *supra*] or by covalently joining to the immunoglobulin coding sequence all or part of the coding sequence for a non-immunoglobulin polypeptide. Such a non-immunoglobulin polypeptide can be substituted for the constant domains of an antibody of the invention, or can be substituted for the variable domains of one antigen-combining site of an antibody of the invention to create a chimeric bivalent antibody.

5           The antibodies may be monovalent antibodies. Methods for preparing monovalent antibodies are well known in the art. For example, one method involves recombinant expression of immunoglobulin light chain and modified heavy chain. The heavy chain is truncated generally at any point in the Fc region so as to prevent heavy chain crosslinking. Alternatively, the relevant cysteine residues are substituted with another amino acid residue or are deleted so as to prevent crosslinking.

10           *In vitro* methods are also suitable for preparing monovalent antibodies. Digestion of antibodies to produce fragments thereof, particularly, Fab fragments, can be accomplished using routine techniques known in the art.

### 3.       Human and Humanized Antibodies

15           The anti-PRO antibodies of the invention may further comprise humanized antibodies or human antibodies. Humanized forms of non-human (e.g., murine) antibodies are chimeric immunoglobulins, immunoglobulin chains or fragments thereof (such as Fv, Fab, Fab', F(ab')<sub>2</sub> or other antigen-binding subsequences of antibodies) which contain minimal sequence derived from non-human immunoglobulin. Humanized antibodies include human immunoglobulins (recipient antibody) in which residues from a  
20       complementary determining region (CDR) of the recipient are replaced by residues from a CDR of a non-human species (donor antibody) such as mouse, rat or rabbit having the desired specificity, affinity and capacity. In some instances, Fv framework residues of the human immunoglobulin are replaced by corresponding non-human residues. Humanized antibodies may also comprise residues which are found neither in the recipient antibody nor in the imported CDR or framework sequences. In general, the humanized antibody will comprise substantially  
25       all of at least one, and typically two, variable domains, in which all or substantially all of the CDR regions correspond to those of a non-human immunoglobulin and all or substantially all of the FR regions are those of a human immunoglobulin consensus sequence. The humanized antibody optimally also will comprise at least a portion of an immunoglobulin constant region (Fc), typically that of a human immunoglobulin [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-329 (1988); and Presta, *Curr. Op. Struct. Biol.*,  
30       2:593-596 (1992)].

          Methods for humanizing non-human antibodies are well known in the art. Generally, a humanized antibody has one or more amino acid residues introduced into it from a source which is non-human. These non-human amino acid residues are often referred to as "import" residues, which are typically taken from an "import" variable domain. Humanization can be essentially performed following the method of Winter and co-workers [Jones et al., *Nature*, 321:522-525 (1986); Riechmann et al., *Nature*, 332:323-327 (1988); Verhoeven et al., *Science*, 239:1534-1536 (1988)], by substituting rodent CDRs or CDR sequences for the corresponding sequences of a human antibody. Accordingly, such "humanized" antibodies are chimeric antibodies (U.S. Patent No.

4,816,567), wherein substantially less than an intact human variable domain has been substituted by the corresponding sequence from a non-human species. In practice, humanized antibodies are typically human antibodies in which some CDR residues and possibly some FR residues are substituted by residues from analogous sites in rodent antibodies.

Human antibodies can also be produced using various techniques known in the art, including phage display libraries [Hoogenboom and Winter, *J. Mol. Biol.*, 227:381 (1991); Marks et al., *J. Mol. Biol.*, 222:581 (1991)]. The techniques of Cole et al. and Boerner et al. are also available for the preparation of human monoclonal antibodies (Cole et al., *Monoclonal Antibodies and Cancer Therapy*, Alan R. Liss, p. 77 (1985) and Boerner et al., *J. Immunol.*, 147(1):86-95 (1991)]. Similarly, human antibodies can be made by introducing of human immunoglobulin loci into transgenic animals, e.g., mice in which the endogenous immunoglobulin genes have been partially or completely inactivated. Upon challenge, human antibody production is observed, which closely resembles that seen in humans in all respects, including gene rearrangement, assembly, and antibody repertoire. This approach is described, for example, in U.S. Patent Nos. 5,545,807; 5,545,806; 5,569,825; 5,625,126; 5,633,425; 5,661,016, and in the following scientific publications: Marks *et al.*, *Bio/Technology* 10, 779-783 (1992); Lonberg *et al.*, *Nature* 368 856-859 (1994); Morrison, *Nature* 368, 812-13 (1994); Fishwild *et al.*, *Nature Biotechnology* 14, 845-51 (1996); Neuberger, *Nature Biotechnology* 14, 826 (1996); Lonberg and Huszar, *Intern. Rev. Immunol.* 13 65-93 (1995).

The antibodies may also be affinity matured using known selection and/or mutagenesis methods as described above. Preferred affinity matured antibodies have an affinity which is five times, more preferably 10 times, even more preferably 20 or 30 times greater than the starting antibody (generally murine, humanized or human) from which the matured antibody is prepared.

#### 4. Bispecific Antibodies

Bispecific antibodies are monoclonal, preferably human or humanized, antibodies that have binding specificities for at least two different antigens. In the present case, one of the binding specificities is for the PRO, the other one is for any other antigen, and preferably for a cell-surface protein or receptor or receptor subunit.

Methods for making bispecific antibodies are known in the art. Traditionally, the recombinant production of bispecific antibodies is based on the co-expression of two immunoglobulin heavy-chain/light-chain pairs, where the two heavy chains have different specificities [Milstein and Cuello, *Nature*, 305:537-539 (1983)]. Because of the random assortment of immunoglobulin heavy and light chains, these hybridomas (quadromas) produce a potential mixture of ten different antibody molecules, of which only one has the correct bispecific structure. The purification of the correct molecule is usually accomplished by affinity chromatography steps. Similar procedures are disclosed in WO 93/08829, published 13 May 1993, and in Traunecker et al., *EMBO J.*, 10:3655-3659 (1991).

Antibody variable domains with the desired binding specificities (antibody-antigen combining sites) can be fused to immunoglobulin constant domain sequences. The fusion preferably is with an immunoglobulin heavy-chain constant domain, comprising at least part of the hinge, CH2, and CH3 regions. It is preferred to have the first heavy-chain constant region (CH1) containing the site necessary for light-chain binding present in at least



one of the fusions. DNAs encoding the immunoglobulin heavy-chain fusions and, if desired, the immunoglobulin light chain, are inserted into separate expression vectors, and are co-transfected into a suitable host organism. For further details of generating bispecific antibodies see, for example, Suresh et al., Methods in Enzymology, 121:210 (1986).

5 According to another approach described in WO 96/27011, the interface between a pair of antibody molecules can be engineered to maximize the percentage of heterodimers which are recovered from recombinant cell culture. The preferred interface comprises at least a part of the CH3 region of an antibody constant domain. In this method, one or more small amino acid side chains from the interface of the first antibody molecule are replaced with larger side chains (e.g. tyrosine or tryptophan). Compensatory "cavities" of identical or similar size to the large side chain(s) are created on the interface of the second antibody molecule by replacing large  
10 amino acid side chains with smaller ones (e.g. alanine or threonine). This provides a mechanism for increasing the yield of the heterodimer over other unwanted end-products such as homodimers.

Bispecific antibodies can be prepared as full length antibodies or antibody fragments (e.g. F(ab')<sub>2</sub> bispecific antibodies). Techniques for generating bispecific antibodies from antibody fragments have been described in the literature. For example, bispecific antibodies can be prepared using chemical  
15 linkage. Brennan et al., Science 229:81 (1985) describe a procedure wherein intact antibodies are proteolytically cleaved to generate F(ab')<sub>2</sub> fragments. These fragments are reduced in the presence of the dithiol complexing agent sodium arsenite to stabilize vicinal dithiols and prevent intermolecular disulfide formation. The Fab' fragments generated are then converted to thionitrobenzoate (TNB) derivatives. One of the Fab'-TNB derivatives is then reconverted to the Fab'-thiol by reduction with mercaptoethylamine and is mixed with an equimolar  
20 amount of the other Fab'-TNB derivative to form the bispecific antibody. The bispecific antibodies produced can be used as agents for the selective immobilization of enzymes.

Fab' fragments may be directly recovered from *E. coli* and chemically coupled to form bispecific antibodies. Shalaby et al., J. Exp. Med. 175:217-225 (1992) describe the production of a fully humanized bispecific antibody F(ab')<sub>2</sub> molecule. Each Fab' fragment was separately secreted from *E. coli* and subjected to  
25 directed chemical coupling *in vitro* to form the bispecific antibody. The bispecific antibody thus formed was able to bind to cells overexpressing the ErbB2 receptor and normal human T cells, as well as trigger the lytic activity of human cytotoxic lymphocytes against human breast tumor targets.

Various technique for making and isolating bispecific antibody fragments directly from recombinant cell culture have also been described. For example, bispecific antibodies have been produced using leucine zippers.  
30 Kostelny et al., J. Immunol. 148(5):1547-1553 (1992). The leucine zipper peptides from the Fos and Jun proteins were linked to the Fab' portions of two different antibodies by gene fusion. The antibody homodimers were reduced at the hinge region to form monomers and then re-oxidized to form the antibody heterodimers. This method can also be utilized for the production of antibody homodimers. The "diabody" technology described by Hollinger et al., Proc. Natl. Acad. Sci. USA 90:6444-6448 (1993) has provided an alternative mechanism for  
35 making bispecific antibody fragments. The fragments comprise a heavy-chain variable domain (V<sub>H</sub>) connected to a light-chain variable domain (V<sub>L</sub>) by a linker which is too short to allow pairing between the two domains on the same chain. Accordingly, the V<sub>H</sub> and V<sub>L</sub> domains of one fragment are forced to pair with the complementary

V<sub>L</sub> and V<sub>H</sub> domains of another fragment, thereby forming two antigen-binding sites. Another strategy for making bispecific antibody fragments by the use of single-chain Fv (sFv) dimers has also been reported. See, Gruber *et al.*, J. Immunol. 152:5368 (1994).

Antibodies with more than two valencies are contemplated. For example, trispecific antibodies can be prepared. Tutt *et al.*, J. Immunol. 147:60 (1991).

5 Exemplary bispecific antibodies may bind to two different epitopes on a given PRO polypeptide herein. Alternatively, an anti-PRO polypeptide arm may be combined with an arm which binds to a triggering molecule on a leukocyte such as a T-cell receptor molecule (e.g. CD2, CD3, CD28, or B7), or Fc receptors for IgG (FcγR), such as FcγRI (CD64), FcγRII (CD32) and FcγRIII (CD16) so as to focus cellular defense mechanisms to the cell expressing the particular PRO polypeptide. Bispecific antibodies may also be used to localize cytotoxic  
10 agents to cells which express a particular PRO polypeptide. These antibodies possess a PRO-binding arm and an arm which binds a cytotoxic agent or a radionuclide chelator, such as EOTUBE, DPTA, DOTA, or TETA. Another bispecific antibody of interest binds the PRO polypeptide and further binds tissue factor (TF).

#### 5. Heteroconjugate Antibodies

15 Heteroconjugate antibodies are also within the scope of the present invention. Heteroconjugate antibodies are composed of two covalently joined antibodies. Such antibodies have, for example, been proposed to target immune system cells to unwanted cells [U.S. Patent No. 4,676,980], and for treatment of HIV infection [WO 91/00360; WO 92/200373; EP 03089]. It is contemplated that the antibodies may be prepared *in vitro* using known methods in synthetic protein chemistry, including those involving crosslinking agents. For example,  
20 immunotoxins may be constructed using a disulfide exchange reaction or by forming a thioether bond. Examples of suitable reagents for this purpose include iminothiolate and methyl-4-mercaptobutyrimidate and those disclosed, for example, in U.S. Patent No. 4,676,980.

#### 6. Effector Function Engineering

25 It may be desirable to modify the antibody of the invention with respect to effector function, so as to enhance, *e.g.*, the effectiveness of the antibody in treating cancer. For example, cysteine residue(s) may be introduced into the Fc region, thereby allowing interchain disulfide bond formation in this region. The homodimeric antibody thus generated may have improved internalization capability and/or increased complement-mediated cell killing and antibody-dependent cellular cytotoxicity (ADCC). See Caron *et al.*, J. Exp Med., 176:  
30 1191-1195 (1992) and Shopes, J. Immunol., 148: 2918-2922 (1992). Homodimeric antibodies with enhanced anti-tumor activity may also be prepared using heterobifunctional cross-linkers as described in Wolff *et al.* Cancer Research, 53: 2560-2565 (1993). Alternatively, an antibody can be engineered that has dual Fc regions and may thereby have enhanced complement lysis and ADCC capabilities. See Stevenson *et al.*, Anti-Cancer Drug Design,  
35 3: 219-230 (1989).

#### 7. Immunoconjugates

The invention also pertains to immunoconjugates comprising an antibody conjugated to a cytotoxic agent

such as a chemotherapeutic agent, toxin (*e.g.*, an enzymatically active toxin of bacterial, fungal, plant, or animal origin, or fragments thereof), or a radioactive isotope (*i.e.*, a radioconjugate).

Chemotherapeutic agents useful in the generation of such immunoconjugates have been described above. Enzymatically active toxins and fragments thereof that can be used include diphtheria A chain, nonbinding active fragments of diphtheria toxin, exotoxin A chain (from *Pseudomonas aeruginosa*), ricin A chain, abrin A chain, modeccin A chain, alpha-sarcin, *Aleurites fordii* proteins, dianthin proteins, *Phytolaca americana* proteins (PAPI, PAPII, and PAP-S), momordica charantia inhibitor, curcin, crotin, sapaonaria officinalis inhibitor, gelonin, mitogellin, restrictocin, phenomycin, enomycin, and the tricothecenes. A variety of radionuclides are available for the production of radioconjugated antibodies. Examples include  $^{212}\text{Bi}$ ,  $^{131}\text{I}$ ,  $^{131}\text{In}$ ,  $^{90}\text{Y}$ , and  $^{186}\text{Re}$ .

Conjugates of the antibody and cytotoxic agent are made using a variety of bifunctional protein-coupling agents such as N-succinimidyl-3-(2-pyridyldithiol) propionate (SPDP), iminothiolane (IT), bifunctional derivatives of imidoesters (such as dimethyl adipimidate HCL), active esters (such as disuccinimidyl suberate), aldehydes (such as glutaraldehyde), bis-azido compounds (such as bis (p-azidobenzoyl) hexanediamine), bis-diazonium derivatives (such as bis-(p-diazoniumbenzoyl)-ethylenediamine), diisocyanates (such as tolyene 2,6-diisocyanate), and bis-active fluorine compounds (such as 1,5-difluoro-2,4-dinitrobenzene). For example, a ricin immunotoxin can be prepared as described in Vitetta *et al.*, Science, **238**: 1098 (1987). Carbon-14-labeled 1-isothiocyanatobenzyl-3-methyldiethylene triaminepentaacetic acid (MX-DTPA) is an exemplary chelating agent for conjugation of radionucleotide to the antibody. See WO94/11026.

In another embodiment, the antibody may be conjugated to a "receptor" (such streptavidin) for utilization in tumor pretargeting wherein the antibody-receptor conjugate is administered to the patient, followed by removal of unbound conjugate from the circulation using a clearing agent and then administration of a "ligand" (*e.g.*, avidin) that is conjugated to a cytotoxic agent (*e.g.*, a radionucleotide).

#### 8. Immunoliposomes

The antibodies disclosed herein may also be formulated as immunoliposomes. Liposomes containing the antibody are prepared by methods known in the art, such as described in Epstein *et al.*, Proc. Natl. Acad. Sci. USA, **82**: 3688 (1985); Hwang *et al.*, Proc. Natl. Acad. Sci. USA, **77**: 4030 (1980); and U.S. Pat. Nos. 4,485,045 and 4,544,545. Liposomes with enhanced circulation time are disclosed in U.S. Patent No. 5,013,556.

Particularly useful liposomes can be generated by the reverse-phase evaporation method with a lipid composition comprising phosphatidylcholine, cholesterol, and PEG-derivatized phosphatidylethanolamine (PEG-PE). Liposomes are extruded through filters of defined pore size to yield liposomes with the desired diameter. Fab' fragments of the antibody of the present invention can be conjugated to the liposomes as described in Martin *et al.*, J. Biol. Chem., **257**: 286-288 (1982) via a disulfide-interchange reaction. A chemotherapeutic agent (such as Doxorubicin) is optionally contained within the liposome. See Gabizon *et al.*, J. National Cancer Inst., **81**(19): 1484 (1989).

#### 9. Pharmaceutical Compositions of Antibodies

Antibodies specifically binding a PRO polypeptide identified herein, as well as other molecules identified

by the screening assays disclosed hereinbefore, can be administered for the treatment of various disorders in the form of pharmaceutical compositions.

If the PRO polypeptide is intracellular and whole antibodies are used as inhibitors, internalizing antibodies are preferred. However, lipofections or liposomes can also be used to deliver the antibody, or an antibody fragment, into cells. Where antibody fragments are used, the smallest inhibitory fragment that specifically binds to the binding domain of the target protein is preferred. For example, based upon the variable-region sequences of an antibody, peptide molecules can be designed that retain the ability to bind the target protein sequence. Such peptides can be synthesized chemically and/or produced by recombinant DNA technology. See, *e.g.*, Marasco *et al.*, Proc. Natl. Acad. Sci. USA, 90: 7889-7893 (1993). The formulation herein may also contain more than one active compound as necessary for the particular indication being treated, preferably those with complementary activities that do not adversely affect each other. Alternatively, or in addition, the composition may comprise an agent that enhances its function, such as, for example, a cytotoxic agent, cytokine, chemotherapeutic agent, or growth-inhibitory agent. Such molecules are suitably present in combination in amounts that are effective for the purpose intended.

The active ingredients may also be entrapped in microcapsules prepared, for example, by coacervation techniques or by interfacial polymerization, for example, hydroxymethylcellulose or gelatin-microcapsules and poly-(methylmethacrylate) microcapsules, respectively, in colloidal drug delivery systems (for example, liposomes, albumin microspheres, microemulsions, nano-particles, and nanocapsules) or in macroemulsions. Such techniques are disclosed in Remington's Pharmaceutical Sciences, *supra*.

The formulations to be used for *in vivo* administration must be sterile. This is readily accomplished by filtration through sterile filtration membranes.

Sustained-release preparations may be prepared. Suitable examples of sustained-release preparations include semipermeable matrices of solid hydrophobic polymers containing the antibody, which matrices are in the form of shaped articles, *e.g.*, films, or microcapsules. Examples of sustained-release matrices include polyesters, hydrogels (for example, poly(2-hydroxyethyl-methacrylate), or poly(vinylalcohol)), polylactides (U.S. Pat. No. 3,773,919), copolymers of L-glutamic acid and  $\gamma$  ethyl-L-glutamate, non-degradable ethylene-vinyl acetate, degradable lactic acid-glycolic acid copolymers such as the LUPRON DEPOT<sup>TM</sup> (injectable microspheres composed of lactic acid-glycolic acid copolymer and leuprolide acetate), and poly-D-(-)-3-hydroxybutyric acid. While polymers such as ethylene-vinyl acetate and lactic acid-glycolic acid enable release of molecules for over 100 days, certain hydrogels release proteins for shorter time periods. When encapsulated antibodies remain in the body for a long time, they may denature or aggregate as a result of exposure to moisture at 37°C, resulting in a loss of biological activity and possible changes in immunogenicity. Rational strategies can be devised for stabilization depending on the mechanism involved. For example, if the aggregation mechanism is discovered to be intermolecular S-S bond formation through thio-disulfide interchange, stabilization may be achieved by modifying sulfhydryl residues, lyophilizing from acidic solutions, controlling moisture content, using appropriate additives, and developing specific polymer matrix compositions.

G. Uses for anti-PRO Antibodies

The anti-PRO antibodies of the invention have various utilities. For example, anti-PRO antibodies may be used in diagnostic assays for PRO, *e.g.*, detecting its expression (and in some cases, differential expression) in specific cells, tissues, or serum. Various diagnostic assay techniques known in the art may be used, such as competitive binding assays, direct or indirect sandwich assays and immunoprecipitation assays conducted in either heterogeneous or homogeneous phases [Zola, Monoclonal Antibodies: A Manual of Techniques, CRC Press, Inc. (1987) pp. 147-158]. The antibodies used in the diagnostic assays can be labeled with a detectable moiety. The detectable moiety should be capable of producing, either directly or indirectly, a detectable signal. For example, the detectable moiety may be a radioisotope, such as  $^3\text{H}$ ,  $^{14}\text{C}$ ,  $^{32}\text{P}$ ,  $^{35}\text{S}$ , or  $^{125}\text{I}$ , a fluorescent or chemiluminescent compound, such as fluorescein isothiocyanate, rhodamine, or luciferin, or an enzyme, such as alkaline phosphatase, beta-galactosidase or horseradish peroxidase. Any method known in the art for conjugating the antibody to the detectable moiety may be employed, including those methods described by Hunter et al., Nature, 144:945 (1962); David et al., Biochemistry, 13:1014 (1974); Pain et al., J. Immunol. Meth., 40:219 (1981); and Nygren, J. Histochem. and Cytochem., 30:407 (1982).

Anti-PRO antibodies also are useful for the affinity purification of PRO from recombinant cell culture or natural sources. In this process, the antibodies against PRO are immobilized on a suitable support, such as Sephadex resin or filter paper, using methods well known in the art. The immobilized antibody then is contacted with a sample containing the PRO to be purified, and thereafter the support is washed with a suitable solvent that will remove substantially all the material in the sample except the PRO, which is bound to the immobilized antibody. Finally, the support is washed with another suitable solvent that will release the PRO from the antibody.

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby incorporated by reference in their entirety.

EXAMPLES

Commercially available reagents referred to in the examples were used according to manufacturer's instructions unless otherwise indicated. The source of those cells identified in the following examples, and throughout the specification, by ATCC accession numbers is the American Type Culture Collection, Manassas, VA.

EXAMPLE 1: Extracellular Domain Homology Screening to Identify Novel Polypeptides and cDNA Encoding Therefor

The extracellular domain (ECD) sequences (including the secretion signal sequence, if any) from about 950 known secreted proteins from the Swiss-Prot public database were used to search EST databases. The EST databases included public databases (*e.g.*, Dayhoff, GenBank), and proprietary databases (*e.g.* LIFESEQ™, Incyte Pharmaceuticals, Palo Alto, CA). The search was performed using the computer program BLAST or

BLAST-2 (Altschul *et al.*, Methods in Enzymology, 266:460-480 (1996)) as a comparison of the ECD protein sequences to a 6 frame translation of the EST sequences. Those comparisons with a BLAST score of 70 (or in some cases 90) or greater that did not encode known proteins were clustered and assembled into consensus DNA sequences with the program "phrap" (Phil Green, University of Washington, Seattle, WA).

Using this extracellular domain homology screen, consensus DNA sequences were assembled relative to the other identified EST sequences using phrap. In addition, the consensus DNA sequences obtained were often (but not always) extended using repeated cycles of BLAST or BLAST-2 and phrap to extend the consensus sequence as far as possible using the sources of EST sequences discussed above.

Based upon the consensus sequences obtained as described above, oligonucleotides were then synthesized and used to identify by PCR a cDNA library that contained the sequence of interest and for use as probes to isolate a clone of the full-length coding sequence for a PRO polypeptide. Forward and reverse PCR primers generally range from 20 to 30 nucleotides and are often designed to give a PCR product of about 100-1000 bp in length. The probe sequences are typically 40-55 bp in length. In some cases, additional oligonucleotides are synthesized when the consensus sequence is greater than about 1-1.5kbp. In order to screen several libraries for a full-length clone, DNA from the libraries was screened by PCR amplification, as per Ausubel *et al.*, Current Protocols in Molecular Biology, with the PCR primer pair. A positive library was then used to isolate clones encoding the gene of interest using the probe oligonucleotide and one of the primer pairs.

The cDNA libraries used to isolate the cDNA clones were constructed by standard methods using commercially available reagents such as those from Invitrogen, San Diego, CA. The cDNA was primed with oligo dT containing a NotI site, linked with blunt to SalI hemikinased adaptors, cleaved with NotI, sized appropriately by gel electrophoresis, and cloned in a defined orientation into a suitable cloning vector (such as pRKB or pRKD; pRK5B is a precursor of pRK5D that does not contain the SfiI site; *see*, Holmes *et al.*, Science, 253:1278-1280 (1991)) in the unique XhoI and NotI sites.

#### EXAMPLE 2: Isolation of cDNA clones by Amylase Screening

##### 1. Preparation of oligo dT primed cDNA library

mRNA was isolated from a human tissue of interest using reagents and protocols from Invitrogen, San Diego, CA (Fast Track 2). This RNA was used to generate an oligo dT primed cDNA library in the vector pRK5D using reagents and protocols from Life Technologies, Gaithersburg, MD (Super Script Plasmid System). In this procedure, the double stranded cDNA was sized to greater than 1000 bp and the SalI/NotI linkered cDNA was cloned into XhoI/NotI cleaved vector. pRK5D is a cloning vector that has an sp6 transcription initiation site followed by an SfiI restriction enzyme site preceding the XhoI/NotI cDNA cloning sites.

##### 2. Preparation of random primed cDNA library

A secondary cDNA library was generated in order to preferentially represent the 5' ends of the primary cDNA clones. Sp6 RNA was generated from the primary library (described above), and this RNA was used to generate a random primed cDNA library in the vector pSST-AMY.0 using reagents and protocols from Life Technologies (Super Script Plasmid System, referenced above). In this procedure the double stranded cDNA was

sized to 500-1000 bp, linked with blunt to NotI adaptors, cleaved with SfiI, and cloned into SfiI/NotI cleaved vector. pSST-AMY.0 is a cloning vector that has a yeast alcohol dehydrogenase promoter preceding the cDNA cloning sites and the mouse amylase sequence (the mature sequence without the secretion signal) followed by the yeast alcohol dehydrogenase terminator, after the cloning sites. Thus, cDNAs cloned into this vector that are fused in frame with amylase sequence will lead to the secretion of amylase from appropriately transfected yeast colonies.

### 3. Transformation and Detection

DNA from the library described in paragraph 2 above was chilled on ice to which was added electrocompetent DH10B bacteria (Life Technologies, 20 ml). The bacteria and vector mixture was then electroporated as recommended by the manufacturer. Subsequently, SOC media (Life Technologies, 1 ml) was added and the mixture was incubated at 37°C for 30 minutes. The transformants were then plated onto 20 standard 150 mm LB plates containing ampicillin and incubated for 16 hours (37°C). Positive colonies were scraped off the plates and the DNA was isolated from the bacterial pellet using standard protocols, *e.g.* CsCl-gradient. The purified DNA was then carried on to the yeast protocols below.

The yeast methods were divided into three categories: (1) Transformation of yeast with the plasmid/cDNA combined vector; (2) Detection and isolation of yeast clones secreting amylase; and (3) PCR amplification of the insert directly from the yeast colony and purification of the DNA for sequencing and further analysis.

The yeast strain used was HD56-5A (ATCC-90785). This strain has the following genotype: MAT alpha, *ura3-52*, *leu2-3*, *leu2-112*, *his3-11*, *his3-15*, *MAL*<sup>+</sup>, *SUC*<sup>+</sup>, *GAL*<sup>+</sup>. Preferably, yeast mutants can be employed that have deficient post-translational pathways. Such mutants may have translocation deficient alleles in *sec71*, *sec72*, *sec62*, with truncated *sec71* being most preferred. Alternatively, antagonists (including antisense nucleotides and/or ligands) which interfere with the normal operation of these genes, other proteins implicated in this post translation pathway (*e.g.*, SEC61p, SEC72p, SEC62p, SEC63p, TDJ1p or SSA1p-4p) or the complex formation of these proteins may also be preferably employed in combination with the amylase-expressing yeast.

Transformation was performed based on the protocol outlined by Gietz *et al.*, Nucl. Acid. Res., 20:1425 (1992). Transformed cells were then inoculated from agar into YEPD complex media broth (100 ml) and grown overnight at 30°C. The YEPD broth was prepared as described in Kaiser *et al.*, Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 207 (1994). The overnight culture was then diluted to about 2 x 10<sup>6</sup> cells/ml (approx. OD<sub>600</sub>=0.1) into fresh YEPD broth (500 ml) and regrown to 1 x 10<sup>7</sup> cells/ml (approx. OD<sub>600</sub>=0.4-0.5).

The cells were then harvested and prepared for transformation by transfer into GS3 rotor bottles in a Sorval GS3 rotor at 5,000 rpm for 5 minutes, the supernatant discarded, and then resuspended into sterile water, and centrifuged again in 50 ml falcon tubes at 3,500 rpm in a Beckman GS-6KR centrifuge. The supernatant was discarded and the cells were subsequently washed with LiAc/TE (10 ml, 10 mM Tris-HCl, 1 mM EDTA pH 7.5, 100 mM Li<sub>2</sub>OOCCH<sub>3</sub>), and resuspended into LiAc/TE (2.5 ml).

Transformation took place by mixing the prepared cells (100 µl) with freshly denatured single stranded

salmon testes DNA (Lofstrand Labs, Gaithersburg, MD) and transforming DNA (1  $\mu$ g, vol. < 10  $\mu$ l) in microfuge tubes. The mixture was mixed briefly by vortexing, then 40% PEG/TE (600  $\mu$ l, 40% polyethylene glycol-4000, 10 mM Tris-HCl, 1 mM EDTA, 100 mM Li<sub>2</sub>OOCCH<sub>3</sub>, pH 7.5) was added. This mixture was gently mixed and incubated at 30°C while agitating for 30 minutes. The cells were then heat shocked at 42°C for 15 minutes, and the reaction vessel centrifuged in a microfuge at 12,000 rpm for 5-10 seconds, decanted and resuspended into TE (500  $\mu$ l, 10 mM Tris-HCl, 1 mM EDTA pH 7.5) followed by recentrifugation. The cells were then diluted into TE (1 ml) and aliquots (200  $\mu$ l) were spread onto the selective media previously prepared in 150 mm growth plates (VWR).

Alternatively, instead of multiple small reactions, the transformation was performed using a single, large scale reaction, wherein reagent amounts were scaled up accordingly.

The selective media used was a synthetic complete dextrose agar lacking uracil (SCD-Ura) prepared as described in Kaiser *et al.*, Methods in Yeast Genetics, Cold Spring Harbor Press, Cold Spring Harbor, NY, p. 208-210 (1994). Transformants were grown at 30°C for 2-3 days.

The detection of colonies secreting amylase was performed by including red starch in the selective growth media. Starch was coupled to the red dye (Reactive Red-120, Sigma) as per the procedure described by Biely *et al.*, Anal. Biochem., 172:176-179 (1988). The coupled starch was incorporated into the SCD-Ura agar plates at a final concentration of 0.15% (w/v), and was buffered with potassium phosphate to a pH of 7.0 (50-100 mM final concentration).

The positive colonies were picked and streaked across fresh selective media (onto 150 mm plates) in order to obtain well isolated and identifiable single colonies. Well isolated single colonies positive for amylase secretion were detected by direct incorporation of red starch into buffered SCD-Ura agar. Positive colonies were determined by their ability to break down starch resulting in a clear halo around the positive colony visualized directly.

#### 4. Isolation of DNA by PCR Amplification

When a positive colony was isolated, a portion of it was picked by a toothpick and diluted into sterile water (30  $\mu$ l) in a 96 well plate. At this time, the positive colonies were either frozen and stored for subsequent analysis or immediately amplified. An aliquot of cells (5  $\mu$ l) was used as a template for the PCR reaction in a 25  $\mu$ l volume containing: 0.5  $\mu$ l KlenTaq (Clontech, Palo Alto, CA); 4.0  $\mu$ l 10 mM dNTP's (Perkin Elmer-Cetus); 2.5  $\mu$ l KlenTaq buffer (Clontech); 0.25  $\mu$ l forward oligo 1; 0.25  $\mu$ l reverse oligo 2; 12.5  $\mu$ l distilled water. The sequence of the forward oligonucleotide 1 was:

5'-TGTAACACGACGGCCAGTTAAATAGACCTGCAATTATTAATCT-3' (SEQ ID NO:611)

The sequence of reverse oligonucleotide 2 was:

5'-CAGGAAACAGCTATGACCACCTGCACACCTGCAAATCCATT-3' (SEQ ID NO:612)

PCR was then performed as follows:

- |    |              |                  |
|----|--------------|------------------|
| a. | Denature     | 92°C, 5 minutes  |
| b. | 3 cycles of: |                  |
|    | Denature     | 92°C, 30 seconds |
|    | Anneal       | 59°C, 30 seconds |



		Extend	72°C, 60 seconds
5	c.	3 cycles of:	
		Denature	92°C, 30 seconds
		Anneal	57°C, 30 seconds
		Extend	72°C, 60 seconds
	d.	25 cycles of:	
		Denature	92°C, 30 seconds
		Anneal	55°C, 30 seconds
		Extend	72°C, 60 seconds
10	e.	Hold	4°C

The underlined regions of the oligonucleotides annealed to the ADH promoter region and the amylase region, respectively, and amplified a 307 bp region from vector pSST-AMY.0 when no insert was present. Typically, the first 18 nucleotides of the 5' end of these oligonucleotides contained annealing sites for the sequencing primers. Thus, the total product of the PCR reaction from an empty vector was 343 bp. However, signal sequence-fused cDNA resulted in considerably longer nucleotide sequences.

Following the PCR, an aliquot of the reaction (5 µl) was examined by agarose gel electrophoresis in a 1% agarose gel using a Tris-Borate-EDTA (TBE) buffering system as described by Sambrook *et al.*, *supra*. Clones resulting in a single strong PCR product larger than 400 bp were further analyzed by DNA sequencing after purification with a 96 Qiaquick PCR clean-up column (Qiagen Inc., Chatsworth, CA).

#### EXAMPLE 3: Isolation of cDNA Clones Using Signal Algorithm Analysis

Various polypeptide-encoding nucleic acid sequences were identified by applying a proprietary signal sequence finding algorithm developed by Genentech, Inc. (South San Francisco, CA) upon ESTs as well as clustered and assembled EST fragments from public (*e.g.*, GenBank) and/or private (LIFESEQ®, Incyte Pharmaceuticals, Inc., Palo Alto, CA) databases. The signal sequence algorithm computes a secretion signal score based on the character of the DNA nucleotides surrounding the first and optionally the second methionine codon(s) (ATG) at the 5'-end of the sequence or sequence fragment under consideration. The nucleotides following the first ATG must code for at least 35 unambiguous amino acids without any stop codons. If the first ATG has the required amino acids, the second is not examined. If neither meets the requirement, the candidate sequence is not scored. In order to determine whether the EST sequence contains an authentic signal sequence, the DNA and corresponding amino acid sequences surrounding the ATG codon are scored using a set of seven sensors (evaluation parameters) known to be associated with secretion signals. Use of this algorithm resulted in the identification of numerous polypeptide-encoding nucleic acid sequences.

#### EXAMPLE 4: Isolation of cDNA clones Encoding Human PRO Polypeptides

Using the techniques described in Examples 1 to 3 above, numerous full-length cDNA clones were identified as encoding PRO polypeptides as disclosed herein. These cDNAs were then deposited under the terms of the Budapest Treaty with the American Type Culture Collection, 10801 University Blvd., Manassas, VA 20110-2209, USA (ATCC) as shown in Table 7 below.

Table 7

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA16435-1208	209930	June 2, 1998
	DNA23318-1211	209787	April 21, 1998
	DNA23322-1393	203400	October 27, 1998
5	DNA23334-1392	209918	June 2, 1998
	DNA26843-1389	203099	August 4, 1998
	DNA 26844-1394	209926	June 2, 1998
	DNA30867-1335	209807	April 28, 1998
	DNA33470-1175	209398	October 17, 1997
10	DNA34436-1238	209523	December 10, 1997
	DNA35557-1137	209255	September 16, 1997
	DNA35599-1168	209373	October 16, 1997
	DNA35668-1171	209371	October 16, 1997
	DNA36992-1168	209382	October 16, 1997
15	DNA39423-1182	209387	October 17, 1997
	DNA39427-1179	209395	October 17, 1997
	DNA39510-1181	209392	October 17, 1997
	DNA39518-1247	209529	December 10, 1997
	DNA39975-1210	209783	April 21, 1998
20	DNA39976-1215	209524	December 10, 1997
	DNA39979-1213	209789	April 21, 1998
	DNA40594-1233	209617	February 5, 1998
	DNA40603-1232	209486	November 21, 1997
	DNA40604-1187	209394	October 17, 1997
25	DNA40625-1189	209788	April 21, 1998
	DNA41225-1217	209491	November 21, 1997
	DNA41379-1236	209488	November 21, 1997
	DNA41386-1316	209703	March 26, 1998
	DNA44161-1434	209907	May 27, 1998
30	DNA44179-1362	209851	May 6, 1998
	DNA44192-1246	209531	December 10, 1997
	DNA44694-1500	203114	August 11, 1998
	DNA45234-1277	209654	March 5, 1998
	DNA45409-2511	203579	January 12, 1999
35	DNA45415-1318	209810	April 28, 1998
	DNA45417-1432	209910	May 27, 1998
	DNA45493-1349	209805	April 28, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA46776-1284	209721	March 31, 1998
	DNA48296-1292	209668	March 11, 1998
	DNA48306-1291	209911	May 27, 1998
5	DNA48328-1355	209843	May 6, 1998
	DNA48329-1290	209785	April 21, 1998
	DNA48334-1435	209924	June 2, 1998
	DNA49141-1431	203003	June 23, 1998
	DNA49624-1279	209655	March 5, 1998
10	DNA49647-1398	209919	June 2, 1998
	DNA49819-1439	209931	June 2, 1998
	DNA50911-1288	209714	March 31, 1998
	DNA50914-1289	209722	March 31, 1998
	DNA50919-1361	209848	May 6, 1998
15	DNA50980-1286	209717	March 31, 1998
	DNA52185-1370	209861	May 14, 1998
	DNA53906-1368	209747	April 7, 1998
	DNA53912-1457	209870	May 14, 1998
	DNA53913-1490	203162	August 25, 1998
20	DNA53977-1371	209862	May 14, 1998
	DNA53978-1443	209983	June 16, 1998
	DNA53996-1442	209921	June 2, 1998
	DNA54002-1367	209754	April 7, 1998
	DNA55737-1345	209753	April 7, 1998
25	DNA56050-1455	203011	June 23, 1998
	DNA56052-1454	203026	June 23, 1998
	DNA56107-1415	203405	October 27, 1998
	DNA56110-1437	203113	August 11, 1998
	DNA56406-1704	203478	November 17, 1998
30	DNA56409-1377	209882	May 20, 1998
	DNA56410-1414	209923	June 2, 1998
	DNA56436-1448	209902	May 27, 1998
	DNA56529-1647	203293	September 29, 1998
	DNA56855-1447	203004	June 23, 1998
35	DNA56859-1445	203019	June 23, 1998
	DNA56860-1510	209952	June 9, 1998
	DNA56865-1491	203022	June 23, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA56868-1478	203024	June 23, 1998
	DNA56869-1545	203161	August 25, 1998
	DNA56870-1492	209925	June 2, 1998
5	DNA57039-1402	209777	April 14, 1998
	DNA57253-1382	209867	May 14, 1998
	DNA57254-1477	203289	September 29, 1998
	DNA57699-1412	203020	June 23, 1998
	DNA57704-1452	209953	June 9, 1998
10	DNA57710-1451	203048	July 1, 1998
	DNA57827-1493	203045	July 1, 1998
	DNA57844-1410	203010	June 23, 1998
	DNA58723-1588	203133	August 18, 1998
	DNA58727-1474	203171	September 1, 1998
15	DNA58730-1607	203221	September 15, 1998
	DNA58732-1650	203290	September 29, 1998
	DNA58737-1473	203136	August 18, 1998
	DNA58743-1609	203154	August 25, 1998
	DNA58747-1384	209868	May 14, 1998
20	DNA58828-1519	203172	September 1, 1998
	DNA58846-1409	209957	June 9, 1998
	DNA58848-1472	209955	June 9, 1998
	DNA58849-1494	209958	June 9, 1998
	DNA58850-1495	209956	June 9, 1998
25	DNA58852-1637	203271	September 22, 1998
	DNA58853-1423	203016	June 23, 1998
	DNA58855-1422	203018	June 23, 1998
	DNA59211-1450	209960	June 9, 1998
	DNA59212-1627	203245	September 9, 1998
30	DNA59213-1487	209959	June 9, 1998
	DNA59219-1613	203220	September 15, 1998
	DNA59497-1496	209941	June 4, 1998
	DNA59602-1436	203051	July 1, 1998
	DNA59603-1419	209944	June 9, 1998
35	DNA59605-1418	203005	June 23, 1998
	DNA59607-1497	209946	June 9, 1998
	DNA59610-1556	209990	June 16, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA59612-1466	209947	June 9, 1998
	DNA59613-1417	203007	June 23, 1998
	DNA59616-1465	209991	June 16, 1998
5	DNA59619-1464	203041	July 1, 1998
	DNA59625-1498	209992	June 16, 1998
	DNA59817-1703	203470	November 17, 1998
	DNA59827-1426	203089	August 4, 1998
	DNA59828-1608	203158	August 25, 1998
10	DNA59837-2545	203658	February 9, 1999
	DNA59844-2542	203650	February 9, 1999
	DNA59853-1505	209985	June 16, 1998
	DNA59854-1459	209974	June 16, 1998
	DNA59855-1485	209987	June 16, 1998
15	DNA60278-1530	203170	September 1, 1998
	DNA60283-1484	203043	July 1, 1998
	DNA60608-1577	203126	August 18, 1998
	DNA60611-1524	203175	September 1, 1998
	DNA60619-1482	209993	June 16, 1998
20	DNA60625-1507	209975	June 16, 1998
	DNA60629-1481	209979	June 16, 1998
	DNA60740-1615	203456	November 3, 1998
	DNA61608-1606	203239	September 9, 1998
	DNA61755-1554	203112	August 11, 1998
25	DNA62809-1531	203237	September 9, 1998
	DNA62812-1594	203248	September 9, 1998
	DNA62813-2544	203655	February 9, 1999
	DNA62845-1684	203361	October 20, 1998
	DNA64849-1604	203468	November 17, 1998
30	DNA64852-1589	203127	August 18, 1998
	DNA64863-1573	203251	September 9, 1998
	DNA64881-1602	203240	September 9, 1998
	DNA64902-1667	203317	October 6, 1998
	DNA64952-1568	203222	September 15, 1998
35	DNA65403-1565	203230	September 15, 1998
	DNA65413-1534	203234	September 15, 1998
	DNA65423-1595	203227	September 15, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA66304-1546	203321	October 6, 1998
	DNA66308-1537	203159	August 25, 1998
	DNA66511-1563	203228	September 15, 1998
5	DNA66512-1564	203218	September 15, 1998
	DNA66519-1535	203236	September 15, 1998
	DNA66521-1583	203225	September 15, 1998
	DNA66658-1584	203229	September 15, 1998
	DNA66660-1585	203279	September 22, 1998
10	DNA66669-1597	203272	September 22, 1998
	DNA66674-1599	203281	September 22, 1998
	DNA68836-1656	203455	November 3, 1998
	DNA68862-2546	203652	February 9, 1999
	DNA68866-1644	203283	September 22, 1998
15	DNA68869-1610	203164	August 25, 1998
	DNA68871-1638	203280	September 22, 1998
	DNA68879-1631	203274	September 22, 1998
	DNA68880-1676	203319	October 6, 1998
	DNA68882-1677	203318	October 6, 1998
20	DNA68883-1691	203535	December 15, 1998
	DNA68885-1678	203311	October 6, 1998
	DNA71180-1655	203403	October 27, 1998
	DNA71184-1634	203266	September 22, 1998
	DNA71213-1659	203401	October 27, 1998
25	DNA71234-1651	203402	October 27, 1998
	DNA71269-1621	203284	September 22, 1998
	DNA71277-1636	203285	September 22, 1998
	DNA71286-1687	203357	October 20, 1998
	DNA71883-1660	203475	November 17, 1998
30	DNA73401-1633	203273	September 22, 1998
	DNA73492-1671	203324	October 6, 1998
	DNA73730-1679	203320	October 6, 1998
	DNA73734-1680	203363	October 20, 1998
	DNA73735-1681	203356	October 20, 1998
35	DNA73742-1662	203316	October 6, 1998
	DNA73746-1654	203411	October 27, 1998
	DNA73760-1672	203314	October 6, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA76393-1664	203323	October 6, 1998
	DNA76398-1699	203474	November 17, 1998
	DNA76399-1700	203472	November 17, 1998
5	DNA76522-2500	203469	November 17, 1998
	DNA76533-1689	203410	October 27, 1998
	DNA77303-2502	203479	November 17, 1998
	DNA77626-1705	203536	December 15, 1998
	DNA77648-1688	203408	October 27, 1998
10	DNA81754-2532	203542	December 15, 1998
	DNA81757-2512	203543	December 15, 1998
	DNA82302-2529	203534	December 15, 1998
	DNA82340-2530	203547	December 22, 1998
	DNA87991-2540	203656	February 9, 1999
15	DNA92238-2539	203602	January 20, 1999
	DNA115291-2681	PTA-202	June 8, 1999
	DNA23336-2861	PTA-1673	April 11, 2000
	DNA30862-1396	209920	June 2, 1998
	DNA30871-1157	209380	October 16, 1997
20	DNA32279-1131	209259	September 16, 1997
	DNA33206-1165	209372	October 16, 1997
	DNA35673-1201	209418	October 28, 1997
	DNA47361-1154-2	209431	November 7, 1997
	DNA49631-1328	209806	April 28, 1998
25	DNA52594-1270	209679	March 17, 1998
	DNA55800-1263	209680	March 17, 1998
	DNA56531-1648	203286	September 29, 1998
	DNA56965-1356	209842	May 6, 1998
	DNA57037-1444	209903	May 27, 1998
30	DNA57695-1340	203006	June 23, 1998
	DNA57834-1339	209954	June 9, 1998
	DNA57841-1522	203458	November 3, 1998
	DNA58847-1383	209879	May 20, 1998
	DNA59493-1420	203050	July 1, 1998
35	DNA59586-1520	203288	September 29, 1998
	DNA59608-2577	203870	March 23, 1999
	DNA59849-1504	209986	June 16, 1998

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA60292-1506	203540	December 15, 1998
	DNA62377-1381-1	203552	December 22, 1998
	DNA62880-1513	203097	August 4, 1998
5	DNA66672-1586	203265	September 22, 1998
	DNA67962-1649	203291	September 29, 1998
	DNA69555-2867	PTA-1632	April 4, 2000
	DNA71162-2764	PTA-860	October 19, 1999
	DNA71290-1630	203275	September 22, 1998
10	DNA76401-1683	203360	October 20, 1998
	DNA76541-1675	203409	October 27, 1998
	DNA76788-2526	203551	December 22, 1998
	DNA77623-2524	203546	December 22, 1998
	DNA80136-2503	203541	December 15, 1998
15	DNA83568-2692	PTA-386	July 20, 1999
	DNA84210-2576	203818	March 2, 1999
	DNA86576-2595	203868	March 23, 1999
	DNA87976-2593	203888	March 30, 1999
	DNA92256-2596	203891	March 30, 1999
20	DNA92289-2598	PTA-131	May 25, 1999
	DNA96850-2705	PTA-479	August 3, 1999
	DNA96855-2629	PTA-18	May 4, 1999
	DNA96857-2636	PTA-17	May 4, 1999
	DNA96860-2700	PTA-478	August 3, 1999
25	DNA96861-2844	PTA-1436	March 2, 2000
	DNA96866-2698	PTA-491	August 3, 1999
	DNA96870-2676	PTA-254	June 22, 1999
	DNA96872-2674	PTA-550	August 17, 1999
	DNA96878-2626	PTA-23	May 4, 1999
30	DNA96879-2619	203967	April 27, 1999
	DNA96889-2641	PTA-119	May 25, 1999
	DNA96893-2621	PTA-12	May 4, 1999
	DNA96897-2688	PTA-379	July 20, 1999
	DNA98564-2643	PTA-125	May 25, 1999
35	DNA107443-2718	PTA-490	August 3, 1999
	DNA107786-2723	PTA-474	August 3, 1999
	DNA108682-2712	PTA-486	August 3, 1999



Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA108684-2761	PTA-653	September 14, 1999
	DNA108701-2749	PTA-554	August 17, 1999
	DNA108720-2717	PTA-511	August 10, 1999
5	DNA108726-2729	PTA-514	August 10, 1999
	DNA108728-2760	PTA-654	September 14, 1999
	DNA108738-2767	PTA-862	October 19, 1999
	DNA108743-2722	PTA-508	August 10, 1999
	DNA108758-2759	PTA-655	September 14, 1999
10	DNA108765-2758	PTA-657	September 14, 1999
	DNA108783-2747	PTA-616	August 31, 1999
	DNA108789-2748	PTA-547	August 17, 1999
	DNA108806-2724	PTA-610	August 31, 1999
	DNA108936-2719	PTA-519	August 10, 1999
15	DNA119510-2771	PTA-947	November 9, 1999
	DNA119517-2778	PTA-951	November 16, 1999
	DNA119535-2756	PTA-613	August 31, 1999
	DNA119537-2777	PTA-956	November 16, 1999
	DNA119714-2851	PTA-1537	March 21, 2000
20	DNA125170-2780	PTA-953	November 16, 1999
	DNA129594-2841	PTA-1481	March 14, 2000
	DNA129793-2857	PTA-1733	April 18, 2000
	DNA130809-2769	PTA-949	November 9, 1999
	DNA131639-2874	PTA-1784	April 25, 2000
25	DNA131649-2855	PTA-1482	March 14, 2000
	DNA131652-2876	PTA-1628	April 4, 2000
	DNA131658-2875	PTA-1671	April 11, 2000
	DNA132162-2770	PTA-950	November 9, 1999
	DNA136110-2763	PTA-652	September 14, 1999
30	DNA139592-2866	PTA-1587	March 28, 2000
	DNA139608-2856	PTA-1581	March 28, 2000
	DNA143292-2848	PTA-1778	April 25, 2000
	DNA144844-2843	PTA-1536	March 21, 2000
	DNA144857-2845	PTA-1589	March 28, 2000
35	DNA145841-2868	PTA-1678	April 11, 2000
	DNA148004-2882	PTA-1779	April 25, 2000
	DNA149893-2873	PTA-1672	April 11, 2000

Table 7 (cont')

	<u>Material</u>	<u>ATCC Dep. No.</u>	<u>Deposit Date</u>
	DNA149930-2884	PTA-1668	April 11, 2000
	DNA150157-2898	PTA-1777	April 25, 2000
	DNA150163-2842	PTA-1533	March 21, 2000
5	DNA153579-2894	PTA-1729	April 18, 2000
	DNA164625-2890	PTA-1535	March 21, 2000
	DNA57838-1337	203014	June 23, 1998
	DNA59777-1480	203111	August 11, 1998
	DNA66675-1587	203282	September 22, 1998
10	DNA76532-1702	203473	November 17, 1998
	DNA105849-2704	PTA-473	August 3, 1999
	DNA83500-2506	203391	October 29, 1998

These deposits were made under the provisions of the Budapest Treaty on the International Recognition of the Deposit of Microorganisms for the Purpose of Patent Procedure and the Regulations thereunder (Budapest Treaty). This assures maintenance of a viable culture of the deposit for 30 years from the date of deposit. The deposits will be made available by ATCC under the terms of the Budapest Treaty, and subject to an agreement between Genentech, Inc. and ATCC, which assures permanent and unrestricted availability of the progeny of the culture of the deposit to the public upon issuance of the pertinent U.S. patent or upon laying open to the public of any U.S. or foreign patent application, whichever comes first, and assures availability of the progeny to one determined by the U.S. Commissioner of Patents and Trademarks to be entitled thereto according to 35 USC § 122 and the Commissioner's rules pursuant thereto (including 37 CFR § 1.14 with particular reference to 886 OG 638).

The assignee of the present application has agreed that if a culture of the materials on deposit should die or be lost or destroyed when cultivated under suitable conditions, the materials will be promptly replaced on notification with another of the same. Availability of the deposited material is not to be construed as a license to practice the invention in contravention of the rights granted under the authority of any government in accordance with its patent laws.

#### 30 EXAMPLE 5: Use of PRO as a hybridization probe

The following method describes use of a nucleotide sequence encoding PRO as a hybridization probe.

DNA comprising the coding sequence of full-length or mature PRO as disclosed herein is employed as a probe to screen for homologous DNAs (such as those encoding naturally-occurring variants of PRO) in human tissue cDNA libraries or human tissue genomic libraries.

35 Hybridization and washing of filters containing either library DNAs is performed under the following high stringency conditions. Hybridization of radiolabeled PRO-derived probe to the filters is performed in a solution of 50% formamide, 5x SSC, 0.1% SDS, 0.1% sodium pyrophosphate, 50 mM sodium phosphate, pH

6.8, 2x Denhardt's solution, and 10% dextran sulfate at 42°C for 20 hours. Washing of the filters is performed in an aqueous solution of 0.1x SSC and 0.1% SDS at 42°C.

DNAs having a desired sequence identity with the DNA encoding full-length native sequence PRO can then be identified using standard techniques known in the art.

#### 5 EXAMPLE 6: Expression of PRO in *E. coli*

This example illustrates preparation of an unglycosylated form of PRO by recombinant expression in *E. coli*.

The DNA sequence encoding PRO is initially amplified using selected PCR primers. The primers should contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector. A variety of expression vectors may be employed. An example of a suitable vector is pBR322 (derived from *E. coli*; see Bolivar et al., Gene, 2:95 (1977)) which contains genes for ampicillin and tetracycline resistance. The vector is digested with restriction enzyme and dephosphorylated. The PCR amplified sequences are then ligated into the vector. The vector will preferably include sequences which encode for an antibiotic resistance gene, a trp promoter, a polyhis leader (including the first six STII codons, polyhis sequence, and enterokinase cleavage site), the PRO coding region, lambda transcriptional terminator, and an argU gene.

The ligation mixture is then used to transform a selected *E. coli* strain using the methods described in Sambrook et al., supra. Transformants are identified by their ability to grow on LB plates and antibiotic resistant colonies are then selected. Plasmid DNA can be isolated and confirmed by restriction analysis and DNA sequencing.

Selected clones can be grown overnight in liquid culture medium such as LB broth supplemented with antibiotics. The overnight culture may subsequently be used to inoculate a larger scale culture. The cells are then grown to a desired optical density, during which the expression promoter is turned on.

After culturing the cells for several more hours, the cells can be harvested by centrifugation. The cell pellet obtained by the centrifugation can be solubilized using various agents known in the art, and the solubilized PRO protein can then be purified using a metal chelating column under conditions that allow tight binding of the protein.

PRO may be expressed in *E. coli* in a poly-His tagged form, using the following procedure. The DNA encoding PRO is initially amplified using selected PCR primers. The primers will contain restriction enzyme sites which correspond to the restriction enzyme sites on the selected expression vector, and other useful sequences providing for efficient and reliable translation initiation, rapid purification on a metal chelation column, and proteolytic removal with enterokinase. The PCR-amplified, poly-His tagged sequences are then ligated into an expression vector, which is used to transform an *E. coli* host based on strain 52 (W3110 fuhA(tonA) lon galE rpoHts(htpRts) clpP(lacIq). Transformants are first grown in LB containing 50 mg/ml carbenicillin at 30°C with shaking until an O.D.<sub>600</sub> of 3-5 is reached. Cultures are then diluted 50-100 fold into CRAP media (prepared by mixing 3.57 g (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub>, 0.71 g sodium citrate•2H<sub>2</sub>O, 1.07 g KCl, 5.36 g Difco yeast extract, 5.36 g Sheffield hycase SF in 500 mL water, as well as 110 mM MPOS, pH 7.3, 0.55% (w/v) glucose and 7 mM MgSO<sub>4</sub>) and grown for approximately 20-30 hours at 30°C with shaking. Samples are removed to verify

expression by SDS-PAGE analysis, and the bulk culture is centrifuged to pellet the cells. Cell pellets are frozen until purification and refolding.

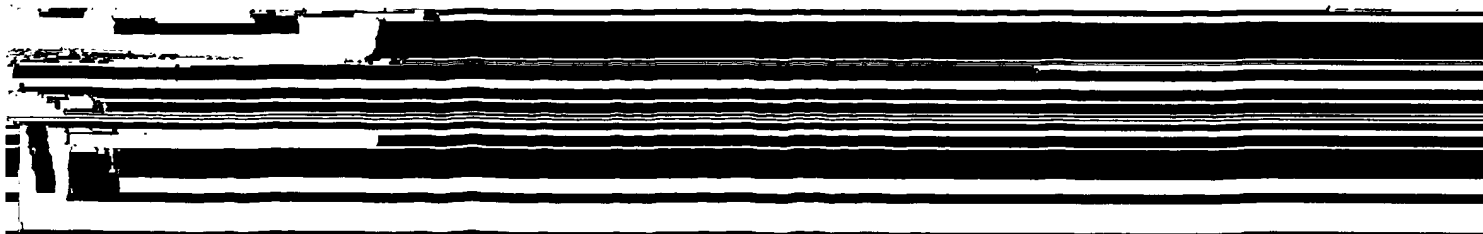
5 *E. coli* paste from 0.5 to 1 L fermentations (6-10 g pellets) is resuspended in 10 volumes (w/v) in 7 M guanidine, 20 mM Tris, pH 8 buffer. Solid sodium sulfite and sodium tetrathionate is added to make final concentrations of 0.1M and 0.02 M, respectively, and the solution is stirred overnight at 4°C. This step results  
10 in a denatured protein with all cysteine residues blocked by sulfitolization. The solution is centrifuged at 40,000 rpm in a Beckman Ultracentrifuge for 30 min. The supernatant is diluted with 3-5 volumes of metal chelate column buffer (6 M guanidine, 20 mM Tris, pH 7.4) and filtered through 0.22 micron filters to clarify. The clarified extract is loaded onto a 5 ml Qiagen Ni-NTA metal chelate column equilibrated in the metal chelate column buffer. The column is washed with additional buffer containing 50 mM imidazole (Calbiochem, Utrol grade), pH 7.4. The protein is eluted with buffer containing 250 mM imidazole. Fractions containing the desired protein are pooled and stored at 4°C. Protein concentration is estimated by its absorbance at 280 nm using the calculated extinction coefficient based on its amino acid sequence.

15 The proteins are refolded by diluting the sample slowly into freshly prepared refolding buffer consisting of: 20 mM Tris, pH 8.6, 0.3 M NaCl, 2.5 M urea, 5 mM cysteine, 20 mM glycine and 1 mM EDTA. Refolding volumes are chosen so that the final protein concentration is between 50 to 100 micrograms/ml. The refolding solution is stirred gently at 4°C for 12-36 hours. The refolding reaction is quenched by the addition of TFA to a final concentration of 0.4% (pH of approximately 3). Before further purification of the protein, the solution is filtered through a 0.22 micron filter and acetonitrile is added to 2-10% final concentration. The refolded protein is chromatographed on a Poros R1/H reversed phase column using a mobile buffer of 0.1% TFA with  
20 elution with a gradient of acetonitrile from 10 to 80%. Aliquots of fractions with A280 absorbance are analyzed on SDS polyacrylamide gels and fractions containing homogeneous refolded protein are pooled. Generally, the properly refolded species of most proteins are eluted at the lowest concentrations of acetonitrile since those species are the most compact with their hydrophobic interiors shielded from interaction with the reversed phase resin. Aggregated species are usually eluted at higher acetonitrile concentrations. In addition to resolving misfolded  
25 forms of proteins from the desired form, the reversed phase step also removes endotoxin from the samples.

Fractions containing the desired folded PRO polypeptide are pooled and the acetonitrile removed using a gentle stream of nitrogen directed at the solution. Proteins are formulated into 20 mM Hepes, pH 6.8 with 0.14 M sodium chloride and 4% mannitol by dialysis or by gel filtration using G25 Superfine (Pharmacia) resins equilibrated in the formulation buffer and sterile filtered.

30 Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### EXAMPLE 7: Expression of PRO in mammalian cells



PRO.

In one embodiment, the selected host cells may be 293 cells. Human 293 cells (ATCC CCL 1573) are grown to confluence in tissue culture plates in medium such as DMEM supplemented with fetal calf serum and optionally, nutrient components and/or antibiotics. About 10  $\mu$ g pRK5-PRO DNA is mixed with about 1  $\mu$ g DNA encoding the VA RNA gene [Thimmappaya et al., *Cell*, 31:543 (1982)] and dissolved in 500  $\mu$ l of 1 mM Tris-HCl, 0.1 mM EDTA, 0.227 M  $\text{CaCl}_2$ . To this mixture is added, dropwise, 500  $\mu$ l of 50 mM HEPES (pH 7.35), 280 mM NaCl, 1.5 mM  $\text{NaPO}_4$ , and a precipitate is allowed to form for 10 minutes at 25°C. The precipitate is suspended and added to the 293 cells and allowed to settle for about four hours at 37°C. The culture medium is aspirated off and 2 ml of 20% glycerol in PBS is added for 30 seconds. The 293 cells are then washed with serum free medium, fresh medium is added and the cells are incubated for about 5 days.

Approximately 24 hours after the transfections, the culture medium is removed and replaced with culture medium (alone) or culture medium containing 200  $\mu$ Ci/ml  $^{35}\text{S}$ -cysteine and 200  $\mu$ Ci/ml  $^{35}\text{S}$ -methionine. After a 12 hour incubation, the conditioned medium is collected, concentrated on a spin filter, and loaded onto a 15% SDS gel. The processed gel may be dried and exposed to film for a selected period of time to reveal the presence of PRO polypeptide. The cultures containing transfected cells may undergo further incubation (in serum free medium) and the medium is tested in selected bioassays.

In an alternative technique, PRO may be introduced into 293 cells transiently using the dextran sulfate method described by Somparyrac et al., *Proc. Natl. Acad. Sci.*, 12:7575 (1981). 293 cells are grown to maximal density in a spinner flask and 700  $\mu$ g pRK5-PRO DNA is added. The cells are first concentrated from the spinner flask by centrifugation and washed with PBS. The DNA-dextran precipitate is incubated on the cell pellet for four hours. The cells are treated with 20% glycerol for 90 seconds, washed with tissue culture medium, and re-introduced into the spinner flask containing tissue culture medium, 5  $\mu$ g/ml bovine insulin and 0.1  $\mu$ g/ml bovine transferrin. After about four days, the conditioned media is centrifuged and filtered to remove cells and debris. The sample containing expressed PRO can then be concentrated and purified by any selected method, such as dialysis and/or column chromatography.

In another embodiment, PRO can be expressed in CHO cells. The pRK5-PRO can be transfected into CHO cells using known reagents such as  $\text{CaPO}_4$  or DEAE-dextran. As described above, the cell cultures can be incubated, and the medium replaced with culture medium (alone) or medium containing a radiolabel such as  $^{35}\text{S}$ -methionine. After determining the presence of PRO polypeptide, the culture medium may be replaced with serum free medium. Preferably, the cultures are incubated for about 6 days, and then the conditioned medium is harvested. The medium containing the expressed PRO can then be concentrated and purified by any selected method.

Epitope-tagged PRO may also be expressed in host CHO cells. The PRO may be subcloned out of the pRK5 vector. The subclone insert can undergo PCR to fuse in frame with a selected epitope tag such as a poly-his tag into a Baculovirus expression vector. The poly-his tagged PRO insert can then be subcloned into a SV40 driven vector containing a selection marker such as DHFR for selection of stable clones. Finally, the CHO cells can be transfected (as described above) with the SV40 driven vector. Labeling may be performed, as described above, to verify expression. The culture medium containing the expressed poly-His tagged PRO can then be

concentrated and purified by any selected method, such as by  $\text{Ni}^{2+}$ -chelate affinity chromatography.

PRO may also be expressed in CHO and/or COS cells by a transient expression procedure or in CHO cells by another stable expression procedure.

Stable expression in CHO cells is performed using the following procedure. The proteins are expressed as an IgG construct (immunoadhesin), in which the coding sequences for the soluble forms (e.g. extracellular domains) of the respective proteins are fused to an IgG1 constant region sequence containing the hinge, CH2 and CH2 domains and/or is a poly-His tagged form.

Following PCR amplification, the respective DNAs are subcloned in a CHO expression vector using standard techniques as described in Ausubel et al., Current Protocols of Molecular Biology, Unit 3.16, John Wiley and Sons (1997). CHO expression vectors are constructed to have compatible restriction sites 5' and 3' of the DNA of interest to allow the convenient shuttling of cDNA's. The vector used expression in CHO cells is as described in Lucas et al., Nucl. Acids Res. 24:9 (1774-1779 (1996), and uses the SV40 early promoter/enhancer to drive expression of the cDNA of interest and dihydrofolate reductase (DHFR). DHFR expression permits selection for stable maintenance of the plasmid following transfection.

Twelve micrograms of the desired plasmid DNA is introduced into approximately 10 million CHO cells using commercially available transfection reagents Superfect<sup>®</sup> (Qiagen), Dosper<sup>®</sup> or Fugene<sup>®</sup> (Boehringer Mannheim). The cells are grown as described in Lucas et al., supra. Approximately  $3 \times 10^7$  cells are frozen in an ampule for further growth and production as described below.

The ampules containing the plasmid DNA are thawed by placement into water bath and mixed by vortexing. The contents are pipetted into a centrifuge tube containing 10 mLs of media and centrifuged at 1000 rpm for 5 minutes. The supernatant is aspirated and the cells are resuspended in 10 mL of selective media (0.2  $\mu\text{m}$  filtered PS20 with 5% 0.2  $\mu\text{m}$  diafiltered fetal bovine serum). The cells are then aliquoted into a 100 mL spinner containing 90 mL of selective media. After 1-2 days, the cells are transferred into a 250 mL spinner filled with 150 mL selective growth medium and incubated at 37°C. After another 2-3 days, 250 mL, 500 mL and 2000 mL spinners are seeded with  $3 \times 10^5$  cells/mL. The cell media is exchanged with fresh media by centrifugation and resuspension in production medium. Although any suitable CHO media may be employed, a production medium described in U.S. Patent No. 5,122,469, issued June 16, 1992 may actually be used. A 3L production spinner is seeded at  $1.2 \times 10^6$  cells/mL. On day 0, the cell number pH is determined. On day 1, the spinner is sampled and sparging with filtered air is commenced. On day 2, the spinner is sampled, the temperature shifted to 33°C, and 30 mL of 500 g/L glucose and 0.6 mL of 10% antifoam (e.g., 35% polydimethylsiloxane emulsion, Dow Corning 365 Medical Grade Emulsion) taken. Throughout the production, the pH is adjusted as necessary to keep it at around 7.2. After 10 days, or until the viability dropped below 70%, the cell culture is harvested by centrifugation and filtering through a 0.22  $\mu\text{m}$  filter. The filtrate was either stored at 4°C or immediately loaded onto columns for purification.

For the poly-His tagged constructs, the proteins are purified using a Ni-NTA column (Qiagen). Before purification, imidazole is added to the conditioned media to a concentration of 5 mM. The conditioned media is pumped onto a 6 ml Ni-NTA column equilibrated in 20 mM Hepes, pH 7.4, buffer containing 0.3 M NaCl and 5 mM imidazole at a flow rate of 4-5 ml/min. at 4°C. After loading, the column is washed with additional

equilibration buffer and the protein eluted with equilibration buffer containing 0.25 M imidazole. The highly purified protein is subsequently desalted into a storage buffer containing 10 mM Hepes, 0.14 M NaCl and 4% mannitol, pH 6.8, with a 25 ml G25 Superfine (Pharmacia) column and stored at -80°C.

Immunoadhesin (Fc-containing) constructs are purified from the conditioned media as follows. The conditioned medium is pumped onto a 5 ml Protein A column (Pharmacia) which had been equilibrated in 20 mM Na phosphate buffer, pH 6.8. After loading, the column is washed extensively with equilibration buffer before elution with 100 mM citric acid, pH 3.5. The eluted protein is immediately neutralized by collecting 1 ml fractions into tubes containing 275  $\mu$ L of 1 M Tris buffer, pH 9. The highly purified protein is subsequently desalted into storage buffer as described above for the poly-His tagged proteins. The homogeneity is assessed by SDS polyacrylamide gels and by N-terminal amino acid sequencing by Edman degradation.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### EXAMPLE 8: Expression of PRO in Yeast

The following method describes recombinant expression of PRO in yeast.

First, yeast expression vectors are constructed for intracellular production or secretion of PRO from the ADH2/GAPDH promoter. DNA encoding PRO and the promoter is inserted into suitable restriction enzyme sites in the selected plasmid to direct intracellular expression of PRO. For secretion, DNA encoding PRO can be cloned into the selected plasmid, together with DNA encoding the ADH2/GAPDH promoter, a native PRO signal peptide or other mammalian signal peptide, or, for example, a yeast alpha-factor or invertase secretory signal/leader sequence, and linker sequences (if needed) for expression of PRO.

Yeast cells, such as yeast strain AB110, can then be transformed with the expression plasmids described above and cultured in selected fermentation media. The transformed yeast supernatants can be analyzed by precipitation with 10% trichloroacetic acid and separation by SDS-PAGE, followed by staining of the gels with Coomassie Blue stain.

Recombinant PRO can subsequently be isolated and purified by removing the yeast cells from the fermentation medium by centrifugation and then concentrating the medium using selected cartridge filters. The concentrate containing PRO may further be purified using selected column chromatography resins.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### EXAMPLE 9: Expression of PRO in Baculovirus-Infected Insect Cells

The following method describes recombinant expression of PRO in Baculovirus-infected insect cells.

The sequence coding for PRO is fused upstream of an epitope tag contained within a baculovirus expression vector. Such epitope tags include poly-his tags and immunoglobulin tags (like Fc regions of IgG). A variety of plasmids may be employed, including plasmids derived from commercially available plasmids such as pVL1393 (Novagen). Briefly, the sequence encoding PRO or the desired portion of the coding sequence of PRO such as the sequence encoding the extracellular domain of a transmembrane protein or the sequence encoding the mature protein if the protein is extracellular is amplified by PCR with primers complementary to the 5' and 3' regions. The 5' primer may incorporate flanking (selected) restriction enzyme sites. The product is then

digested with those selected restriction enzymes and subcloned into the expression vector.

Recombinant baculovirus is generated by co-transfecting the above plasmid and BaculoGold™ virus DNA (Pharmingen) into *Spodoptera frugiperda* ("Sf9") cells (ATCC CRL 1711) using lipofectin (commercially available from GIBCO-BRL). After 4 - 5 days of incubation at 28°C, the released viruses are harvested and used for further amplifications. Viral infection and protein expression are performed as described by O'Reilley et al.,  
5 Baculovirus expression vectors: A Laboratory Manual, Oxford: Oxford University Press (1994).

Expressed poly-his tagged PRO can then be purified, for example, by Ni<sup>2+</sup>-chelate affinity chromatography as follows. Extracts are prepared from recombinant virus-infected Sf9 cells as described by Rupert et al., Nature, 362:175-179 (1993). Briefly, Sf9 cells are washed, resuspended in sonication buffer (25 mL Hepes, pH 7.9; 12.5 mM MgCl<sub>2</sub>; 0.1 mM EDTA; 10% glycerol; 0.1% NP-40; 0.4 M KCl), and sonicated  
10 twice for 20 seconds on ice. The sonicates are cleared by centrifugation, and the supernatant is diluted 50-fold in loading buffer (50 mM phosphate, 300 mM NaCl, 10% glycerol, pH 7.8) and filtered through a 0.45 µm filter. A Ni<sup>2+</sup>-NTA agarose column (commercially available from Qiagen) is prepared with a bed volume of 5 mL, washed with 25 mL of water and equilibrated with 25 mL of loading buffer. The filtered cell extract is loaded onto the column at 0.5 mL per minute. The column is washed to baseline A<sub>280</sub> with loading buffer, at which point  
15 fraction collection is started. Next, the column is washed with a secondary wash buffer (50 mM phosphate; 300 mM NaCl, 10% glycerol, pH 6.0), which elutes nonspecifically bound protein. After reaching A<sub>280</sub> baseline again, the column is developed with a 0 to 500 mM Imidazole gradient in the secondary wash buffer. One mL fractions are collected and analyzed by SDS-PAGE and silver staining or Western blot with Ni<sup>2+</sup>-NTA-conjugated to alkaline phosphatase (Qiagen). Fractions containing the eluted His<sub>10</sub>-tagged PRO are pooled and dialyzed  
20 against loading buffer.

Alternatively, purification of the IgG tagged (or Fc tagged) PRO can be performed using known chromatography techniques, including for instance, Protein A or protein G column chromatography.

Many of the PRO polypeptides disclosed herein were successfully expressed as described above.

#### 25 EXAMPLE 10: Preparation of Antibodies that Bind PRO

This example illustrates preparation of monoclonal antibodies which can specifically bind PRO.

Techniques for producing the monoclonal antibodies are known in the art and are described, for instance, in Goding, supra. Immunogens that may be employed include purified PRO, fusion proteins containing PRO, and cells expressing recombinant PRO on the cell surface. Selection of the immunogen can be made by the skilled  
30 artisan without undue experimentation.

Mice, such as Balb/c, are immunized with the PRO immunogen emulsified in complete Freund's adjuvant and injected subcutaneously or intraperitoneally in an amount from 1-100 micrograms. Alternatively, the immunogen is emulsified in MPL-TDM adjuvant (Ribi Immunochemical Research, Hamilton, MT) and injected into the animal's hind foot pads. The immunized mice are then boosted 10 to 12 days later with additional  
35 immunogen emulsified in the selected adjuvant. Thereafter, for several weeks, the mice may also be boosted with additional immunization injections. Serum samples may be periodically obtained from the mice by retro-orbital bleeding for testing in ELISA assays to detect anti-PRO antibodies.



After a suitable antibody titer has been detected, the animals "positive" for antibodies can be injected

na cell  
which  
idine)

5 are harvested. The spleen cells are then fused (using 35 % polyethylene glycol) to a selected murine myeloma cell line such as P3X63AgU.1, available from ATCC, No. CRL 1597. The fusions generate hybridoma cells which can then be plated in 96 well tissue culture plates containing HAT (hypoxanthine, aminopterin, and thymine) medium to inhibit proliferation of non-fused cells, myeloma hybrids, and spleen cell hybrids.

ion of  
he art.

The hybridoma cells will be screened in an ELISA for reactivity against PRO. Determination of the "positive" hybridoma cells secreting the desired monoclonal antibodies against PRO is within the skill of one of ordinary skill in the art.

roduce  
tissue  
can be  
tively,

10 The positive hybridoma cells can be injected intraperitoneally into syngeneic Balb/c mice to produce ascites containing the anti-PRO monoclonal antibodies. Alternatively, the hybridoma cells can be grown in culture flasks or roller bottles. Purification of the monoclonal antibodies produced in the ascites can be accomplished using ammonium sulfate precipitation, followed by gel exclusion chromatography. Alternatively, affinity chromatography based upon binding of antibody to protein A or protein G can be employed.

#### EXAMPLE 11: Purification of PRO Polypeptides Using Specific Antibodies

the art  
eptide  
st. In  
ody to

15 Native or recombinant PRO polypeptides may be purified by a variety of standard techniques in the art of protein purification. For example, pro-PRO polypeptide, mature PRO polypeptide, or pre-PRO polypeptide may be purified by immunoaffinity chromatography using antibodies specific for the PRO polypeptide of interest. In general, an immunoaffinity column is constructed by covalently coupling the anti-PRO polypeptide antibody to an activated chromatographic resin.

onium  
ewise,  
on or

20 Polyclonal immunoglobulins are prepared from immune sera either by precipitation with ammonium sulfate or by purification on immobilized Protein A (Pharmacia LKB Biotechnology, Piscataway, N.J.). Likewise, monoclonal antibodies are prepared from mouse ascites fluid by ammonium sulfate precipitation.

EXAMPLE 12: Drug Screening

This invention is particularly useful for screening compounds by using PRO polypeptides or binding fragment thereof in any of a variety of drug screening techniques. The PRO polypeptide or fragment employed in such a test may either be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. One method of drug screening utilizes eukaryotic or prokaryotic host cells which are stably transformed with recombinant nucleic acids expressing the PRO polypeptide or fragment. Drugs are screened against such transformed cells in competitive binding assays. Such cells, either in viable or fixed form, can be used for standard binding assays. One may measure, for example, the formation of complexes between PRO polypeptide or a fragment and the agent being tested. Alternatively, one can examine the diminution in complex formation between the PRO polypeptide and its target cell or target receptors caused by the agent being tested.

Thus, the present invention provides methods of screening for drugs or any other agents which can affect a PRO polypeptide-associated disease or disorder. These methods comprise contacting such an agent with an PRO polypeptide or fragment thereof and assaying (i) for the presence of a complex between the agent and the PRO polypeptide or fragment, or (ii) for the presence of a complex between the PRO polypeptide or fragment and the cell, by methods well known in the art. In such competitive binding assays, the PRO polypeptide or fragment is typically labeled. After suitable incubation, free PRO polypeptide or fragment is separated from that present in bound form, and the amount of free or uncomplexed label is a measure of the ability of the particular agent to bind to PRO polypeptide or to interfere with the PRO polypeptide/cell complex.

Another technique for drug screening provides high throughput screening for compounds having suitable binding affinity to a polypeptide and is described in detail in WO 84/03564, published on September 13, 1984. Briefly stated, large numbers of different small peptide test compounds are synthesized on a solid substrate, such as plastic pins or some other surface. As applied to a PRO polypeptide, the peptide test compounds are reacted with PRO polypeptide and washed. Bound PRO polypeptide is detected by methods well known in the art. Purified PRO polypeptide can also be coated directly onto plates for use in the aforementioned drug screening techniques. In addition, non-neutralizing antibodies can be used to capture the peptide and immobilize it on the solid support.

This invention also contemplates the use of competitive drug screening assays in which neutralizing antibodies capable of binding PRO polypeptide specifically compete with a test compound for binding to PRO polypeptide or fragments thereof. In this manner, the antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with PRO polypeptide.

EXAMPLE 13: Rational Drug Design

The goal of rational drug design is to produce structural analogs of biologically active polypeptide of interest (*i.e.*, a PRO polypeptide) or of small molecules with which they interact, *e.g.*, agonists, antagonists, or inhibitors. Any of these examples can be used to fashion drugs which are more active or stable forms of the PRO polypeptide or which enhance or interfere with the function of the PRO polypeptide *in vivo* (*c.f.*, Hodgson, Bio/Technology, 9: 19-21 (1991)).

In one approach, the three-dimensional structure of the PRO polypeptide, or of an PRO

polypeptide-inhibitor complex, is determined by x-ray crystallography, by computer modeling or, most typically, by a combination of the two approaches. Both the shape and charges of the PRO polypeptide must be ascertained to elucidate the structure and to determine active site(s) of the molecule. Less often, useful information regarding the structure of the PRO polypeptide may be gained by modeling based on the structure of homologous proteins. In both cases, relevant structural information is used to design analogous PRO polypeptide-like molecules or to identify efficient inhibitors. Useful examples of rational drug design may include molecules which have improved activity or stability as shown by Braxton and Wells, *Biochemistry*, **31**:7796-7801 (1992) or which act as inhibitors, agonists, or antagonists of native peptides as shown by Athauda *et al.*, *J. Biochem.*, **113**:742-746 (1993).

It is also possible to isolate a target-specific antibody, selected by functional assay, as described above, and then to solve its crystal structure. This approach, in principle, yields a pharmacore upon which subsequent drug design can be based. It is possible to bypass protein crystallography altogether by generating anti-idiotypic antibodies (anti-ids) to a functional, pharmacologically active antibody. As a mirror image of a mirror image, the binding site of the anti-ids would be expected to be an analog of the original receptor. The anti-id could then be used to identify and isolate peptides from banks of chemically or biologically produced peptides. The isolated peptides would then act as the pharmacore.

By virtue of the present invention, sufficient amounts of the PRO polypeptide may be made available to perform such analytical studies as X-ray crystallography. In addition, knowledge of the PRO polypeptide amino acid sequence provided herein will provide guidance to those employing computer modeling techniques in place of or in addition to x-ray crystallography.

#### EXAMPLE 14: Identification of PRO Polypeptides That Stimulate TNF- $\alpha$ Release In Human Blood (Assay 128)

This assay shows that certain PRO polypeptides of the present invention act to stimulate the release of TNF- $\alpha$  in human blood. PRO polypeptides testing positive in this assay are useful for, among other things, research purposes where stimulation of the release of TNF- $\alpha$  would be desired and for the therapeutic treatment of conditions wherein enhanced TNF- $\alpha$  release would be beneficial. Specifically, 200  $\mu$ l of human blood supplemented with 50mM Hepes buffer (pH 7.2) is aliquoted per well in a 96 well test plate. To each well is then added 300 $\mu$ l of either the test PRO polypeptide in 50 mM Hepes buffer (at various concentrations) or 50 mM Hepes buffer alone (negative control) and the plates are incubated at 37°C for 6 hours. The samples are then centrifuged and 50 $\mu$ l of plasma is collected from each well and tested for the presence of TNF- $\alpha$  by ELISA assay. A positive in the assay is a higher amount of TNF- $\alpha$  in the PRO polypeptide treated samples as compared to the negative control samples.

The following PRO polypeptides tested positive in this assay:

PRO1079, PRO827, PRO791, PRO1131, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567, and PRO4333.

#### EXAMPLE 15: Promotion of Chondrocyte Redifferentiation (Assay 129)

This assay is designed to determine whether PRO polypeptides of the present invention show the ability to induce the proliferation and/or redifferentiation of chondrocytes in culture. PRO polypeptides testing positive in this assay would be expected to be useful for the therapeutic treatment of various bone and/or cartilage

disorders such as, for example, sports injuries and arthritis.

Porcine chondrocytes are isolated by overnight collagenase digestion of articular cartilage of the metacarpophalangeal joint of 4-6 month old female pigs. The isolated cells are then seeded at 25,000 cells/cm<sup>2</sup> in Ham F-12 containing 10% FBS and 4 µg/ml gentamycin. The culture media is changed every third day. On day 12, the cells are seeded in 96 well plates at 5,000 cells/well in 100 µl of the same media without serum and 100 µl of either serum-free medium (negative control), staurosporin (final concentration of 5 nM; positive control) or the test PRO polypeptide are added to give a final volume of 200 µl/well. After 5 days at 37°C, 22 µl of media containing 100 µg/ml Hoechst 33342 and 50 µg/ml 5-CFDA is added to each well and incubated for an additional 10 minutes at 37°C. A picture of the green fluorescence is taken for each well and the differentiation state of the chondrocytes is calculated by morphometric analysis. A positive result in the assay is obtained when the >50% of the PRO polypeptide treated cells are differentiated (compared to the background obtained by the negative control).

PRO6029 polypeptide tested positive in this assay.

**EXAMPLE 16: Microarray Analysis to Detect Overexpression of PRO Polypeptides in Cancerous Tumors**

Nucleic acid microarrays, often containing thousands of gene sequences, are useful for identifying differentially expressed genes in diseased tissues as compared to their normal counterparts. Using nucleic acid microarrays, test and control mRNA samples from test and control tissue samples are reverse transcribed and labeled to generate cDNA probes. The cDNA probes are then hybridized to an array of nucleic acids immobilized on a solid support. The array is configured such that the sequence and position of each member of the array is known. For example, a selection of genes known to be expressed in certain disease states may be arrayed on a solid support. Hybridization of a labeled probe with a particular array member indicates that the sample from which the probe was derived expresses that gene. If the hybridization signal of a probe from a test (disease tissue) sample is greater than hybridization signal of a probe from a control (normal tissue) sample, the gene or genes overexpressed in the disease tissue are identified. The implication of this result is that an overexpressed protein in a diseased tissue is useful not only as a diagnostic marker for the presence of the disease condition, but also as a therapeutic target for treatment of the disease condition.

The methodology of hybridization of nucleic acids and microarray technology is well known in the art. In the present example, the specific preparation of nucleic acids for hybridization and probes, slides, and hybridization conditions are all detailed in U.S. Provisional Patent Application Serial No. 60/193,767, filed on March 31, 2000 and which is herein incorporated by reference.

In the present example, cancerous tumors derived from various human tissues were studied for PRO polypeptide-encoding gene expression relative to non-cancerous human tissue in an attempt to identify those PRO polypeptides which are overexpressed in cancerous tumors. Two sets of experimental data were generated. In one set, cancerous human colon tumor tissue and matched non-cancerous human colon tumor tissue from the same patient ("matched colon control") were obtained and analyzed for PRO polypeptide expression using the above described microarray technology. In the second set of data, cancerous human tumor tissue from any of a variety of different human tumors was obtained and compared to a "universal" epithelial control sample which was

prepared by pooling non-cancerous human tissues of epithelial origin, including liver, kidney, and lung. mRNA isolated from the pooled tissues represents a mixture of expressed gene products from these different tissues. Microarray hybridization experiments using the pooled control samples generated a linear plot in a 2-color analysis. The slope of the line generated in a 2-color analysis was then used to normalize the ratios of (test:control detection) within each experiment. The normalized ratios from various experiments were then compared and used

5 to identify clustering of gene expression. Thus, the pooled "universal control" sample not only allowed effective relative gene expression determinations in a simple 2-sample comparison, it also allowed multi-sample comparisons across several experiments.

In the present experiments, nucleic acid probes derived from the herein described PRO polypeptide-encoding nucleic acid sequences were used in the creation of the microarray and RNA from the tumor tissues

above this cutoff ratio were determined to be , demonstrating that various PRO polypeptides s human tumor tissues as compared to a non-demonstrate that the PRO polypeptides of the e presence of one or more cancerous tumors, nors.

was designated as a "cutoff ratio". Only values that were at significant. Table 8 below shows the results of these experiments. of the preent invention are significantly overexpressed in variou cancerous human tissue control. As described above, these data present invention are useful not only as diagnostic markers for th but also serve as therapeutic targets for the treatment of those tur

Table 8

<u>as compared to:</u>	<u>Molecule</u>	<u>is overexpressed in:</u>
universal normal control	PRO276	lung tumor
universal normal control	20 PRO284	colon tumor
universal normal control	PRO284	lung tumor
universal normal control	PRO284	breast tumor
universal normal control	PRO193	colon tumor
universal normal control	PRO193	lung tumor
universal normal control	25 PRO193	breast tumor
univ. normal control	PRO193	prostate tumor

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO234	colon tumor	universal normal control
	PRO234	lung tumor	universal normal control
	PRO234	breast tumor	universal normal control
5	PRO234	liver tumor	universal normal control
	PRO236	colon tumor	universal normal control
	PRO236	lung tumor	universal normal control
	PRO236	breast tumor	universal normal control
	PRO244	colon tumor	universal normal control
10	PRO244	lung tumor	universal normal control
	PRO262	colon tumor	universal normal control
	PRO262	lung tumor	universal normal control
	PRO262	breast tumor	universal normal control
	PRO271	colon tumor	universal normal control
15	PRO271	lung tumor	universal normal control
	PRO268	colon tumor	universal normal control
	PRO268	lung tumor	universal normal control
	PRO268	breast tumor	universal normal control
	PRO270	colon tumor	universal normal control
20	PRO270	lung tumor	universal normal control
	PRO270	breast tumor	universal normal control
	PRO270	liver tumor	universal normal control
	PRO355	lung tumor	universal normal control
	PRO355	breast tumor	universal normal control
25	PRO355	prostate tumor	universal normal control
	PRO298	colon tumor	universal normal control
	PRO298	lung tumor	universal normal control
	PRO298	breast tumor	universal normal control
	PRO299	colon tumor	universal normal control
30	PRO299	lung tumor	universal normal control
	PRO299	breast tumor	universal normal control
	PRO296	colon tumor	universal normal control
	PRO296	breast tumor	universal normal control
	PRO329	colon tumor	universal normal control
35	PRO329	lung tumor	universal normal control
	PRO329	breast tumor	universal normal control
	PRO330	colon tumor	universal normal control
	PRO330	lung tumor	universal normal control
	PRO294	lung tumor	universal normal control
40	PRO294	breast tumor	universal normal control
	PRO300	colon tumor	universal normal control
	PRO300	lung tumor	universal normal control
	PRO300	breast tumor	universal normal control
	PRO307	lung tumor	universal normal control
45	PRO334	colon tumor	universal normal control
	PRO334	lung tumor	universal normal control
	PRO334	breast tumor	universal normal control
	PRO334	prostate tumor	universal normal control
	PRO352	colon tumor	universal normal control
50	PRO352	lung tumor	universal normal control
	PRO352	breast tumor	universal normal control
	PRO352	liver tumor	universal normal control
	PRO710	breast tumor	universal normal control
	PRO873	colon tumor	universal normal control
55	PRO873	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO873	breast tumor	universal normal control
	PRO873	prostate tumor	universal normal control
	PRO354	colon tumor	universal normal control
5	PRO354	lung tumor	universal normal control
	PRO354	breast tumor	universal normal control
	PRO1151	lung tumor	universal normal control
	PRO1151	breast tumor	universal normal control
	PRO382	colon tumor	universal normal control
10	PRO382	lung tumor	universal normal control
	PRO382	breast tumor	universal normal control
	PRO1864	lung tumor	universal normal control
	PRO1864	breast tumor	universal normal control
	PRO1864	liver tumor	universal normal control
15	PRO386	colon tumor	universal normal control
	PRO386	lung tumor	universal normal control
	PRO386	prostate tumor	universal normal control
	PRO541	colon tumor	universal normal control
	PRO541	lung tumor	universal normal control
20	PRO541	breast tumor	universal normal control
	PRO852	breast tumor	universal normal control
	PRO700	colon tumor	universal normal control
	PRO700	lung tumor	universal normal control
	PRO700	breast tumor	universal normal control
25	PRO700	rectal tumor	universal normal control
	PRO708	colon tumor	universal normal control
	PRO708	lung tumor	universal normal control
	PRO708	breast tumor	universal normal control
	PRO707	colon tumor	universal normal control
30	PRO707	lung tumor	universal normal control
	PRO864	colon tumor	universal normal control
	PRO864	lung tumor	universal normal control
	PRO864	breast tumor	universal normal control
	PRO706	colon tumor	universal normal control
35	PRO706	lung tumor	universal normal control
	PRO706	breast tumor	universal normal control
	PRO706	liver tumor	universal normal control
	PRO732	lung tumor	universal normal control
	PRO732	breast tumor	universal normal control
40	PRO732	cervical tumor	universal normal control
	PRO537	colon tumor	universal normal control
	PRO537	lung tumor	universal normal control
	PRO537	breast tumor	universal normal control
	PRO545	lung tumor	universal normal control
45	PRO545	breast tumor	universal normal control
	PRO718	lung tumor	universal normal control
	PRO718	breast tumor	universal normal control
	PRO872	lung tumor	universal normal control
	PRO872	breast tumor	universal normal control
50	PRO872	liver tumor	universal normal control
	PRO704	colon tumor	universal normal control
	PRO704	lung tumor	universal normal control
	PRO704	breast tumor	universal normal control
	PRO705	lung tumor	universal normal control
55	PRO705	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO871	lung tumor	universal normal control
	PRO871	breast tumor	universal normal control
	PRO871	liver tumor	universal normal control
5	PRO702	lung tumor	universal normal control
	PRO944	colon tumor	universal normal control
	PRO944	lung tumor	universal normal control
	PRO944	rectal tumor	universal normal control
	PRO739	lung tumor	universal normal control
10	PRO739	breast tumor	universal normal control
	PRO739	prostate tumor	universal normal control
	PRO941	colon tumor	universal normal control
	PRO941	lung tumor	universal normal control
	PRO941	breast tumor	universal normal control
15	PRO941	rectal tumor	universal normal control
	PRO1082	lung tumor	universal normal control
	PRO1082	breast tumor	universal normal control
	PRO1133	colon tumor	universal normal control
	PRO1133	lung tumor	universal normal control
20	PRO983	colon tumor	universal normal control
	PRO983	lung tumor	universal normal control
	PRO983	breast tumor	universal normal control
	PRO784	colon tumor	universal normal control
	PRO784	lung tumor	universal normal control
25	PRO784	breast tumor	universal normal control
	PRO784	prostate tumor	universal normal control
	PRO783	colon tumor	universal normal control
	PRO783	lung tumor	universal normal control
	PRO783	breast tumor	universal normal control
30	PRO783	liver tumor	universal normal control
	PRO940	colon tumor	universal normal control
	PRO940	lung tumor	universal normal control
	PRO940	breast tumor	universal normal control
	PRO768	colon tumor	universal normal control
35	PRO768	lung tumor	universal normal control
	PRO768	breast tumor	universal normal control
	PRO1079	colon tumor	universal normal control
	PRO1079	lung tumor	universal normal control
	PRO1079	breast tumor	universal normal control
40	PRO1079	rectal tumor	universal normal control
	PRO1078	colon tumor	universal normal control
	PRO1078	lung tumor	universal normal control
	PRO1018	colon tumor	universal normal control
	PRO1018	lung tumor	universal normal control
45	PRO1018	breast tumor	universal normal control
	PRO793	colon tumor	universal normal control
	PRO793	lung tumor	universal normal control
	PRO793	breast tumor	universal normal control
	PRO793	rectal tumor	universal normal control
50	PRO1773	colon tumor	universal normal control
	PRO1773	lung tumor	universal normal control
	PRO1773	prostate tumor	universal normal control
	PRO1014	lung tumor	universal normal control
	PRO1014	breast tumor	universal normal control
55	PRO1013	colon tumor	universal normal control



Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1013	lung tumor	universal normal control
	PRO1013	breast tumor	universal normal control
	PRO1013	liver tumor	universal normal control
5	PRO937	colon tumor	universal normal control
	PRO937	lung tumor	universal normal control
	PRO937	breast tumor	universal normal control
	PRO937	cervical tumor	universal normal control
	PRO937	rectal tumor	universal normal control
10	PRO1477	lung tumor	universal normal control
	PRO1477	breast tumor	universal normal control
	PRO1477	rectal tumor	universal normal control
	PRO842	colon tumor	universal normal control
	PRO842	lung tumor	universal normal control
15	PRO842	breast tumor	universal normal control
	PRO839	colon tumor	universal normal control
	PRO1180	colon tumor	universal normal control
	PRO1180	lung tumor	universal normal control
	PRO1180	liver tumor	universal normal control
20	PRO1134	lung tumor	universal normal control
	PRO1134	breast tumor	universal normal control
	PRO1134	prostate tumor	universal normal control
	PRO1115	colon tumor	universal normal control
	PRO1115	lung tumor	universal normal control
25	PRO1115	breast tumor	universal normal control
	PRO1277	colon tumor	universal normal control
	PRO1277	lung tumor	universal normal control
	PRO1135	lung tumor	universal normal control
	PRO1135	breast tumor	universal normal control
30	PRO1135	cervical tumor	universal normal control
	PRO827	colon tumor	universal normal control
	PRO827	lung tumor	universal normal control
	PRO827	prostate tumor	universal normal control
	PRO827	cervical tumor	universal normal control
35	PRO1057	lung tumor	universal normal control
	PRO1057	breast tumor	universal normal control
	PRO1113	colon tumor	universal normal control
	PRO1113	lung tumor	universal normal control
	PRO1006	colon tumor	universal normal control
40	PRO1006	lung tumor	universal normal control
	PRO1006	breast tumor	universal normal control
	PRO1006	rectal tumor	universal normal control
	PRO1074	lung tumor	universal normal control
	PRO1074	rectal tumor	universal normal control
45	PRO1073	lung tumor	universal normal control
	PRO1073	breast tumor	universal normal control
	PRO1136	colon tumor	universal normal control
	PRO1136	lung tumor	universal normal control
	PRO1136	breast tumor	universal normal control
50	PRO1004	lung tumor	universal normal control
	PRO1344	colon tumor	universal normal control
	PRO1344	lung tumor	universal normal control
	PRO1344	breast tumor	universal normal control
	PRO1344	rectal tumor	universal normal control
55	PRO1110	colon tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1110	lung tumor	universal normal control
	PRO1110	breast tumor	universal normal control
	PRO1378	colon tumor	universal normal control
5	PRO1378	lung tumor	universal normal control
	PRO1378	prostate tumor	universal normal control
	PRO1378	cervical tumor	universal normal control
	PRO1481	colon tumor	universal normal control
	PRO1481	lung tumor	universal normal control
10	PRO1109	lung tumor	universal normal control
	PRO1109	breast tumor	universal normal control
	PRO1383	colon tumor	universal normal control
	PRO1383	lung tumor	universal normal control
	PRO1383	breast tumor	universal normal control
15	PRO1072	lung tumor	universal normal control
	PRO1189	colon tumor	universal normal control
	PRO1189	lung tumor	universal normal control
	PRO1189	breast tumor	universal normal control
	PRO1189	prostate tumor	universal normal control
20	PRO1003	colon tumor	universal normal control
	PRO1003	lung tumor	universal normal control
	PRO1003	breast tumor	universal normal control
	PRO1003	liver tumor	universal normal control
	PRO1003	rectal tumor	universal normal control
25	PRO1108	colon tumor	universal normal control
	PRO1108	lung tumor	universal normal control
	PRO1108	breast tumor	universal normal control
	PRO1137	colon tumor	universal normal control
	PRO1137	lung tumor	universal normal control
30	PRO1137	breast tumor	universal normal control
	PRO1138	colon tumor	universal normal control
	PRO1138	lung tumor	universal normal control
	PRO1138	breast tumor	universal normal control
	PRO1415	colon tumor	universal normal control
35	PRO1415	lung tumor	universal normal control
	PRO1415	prostate tumor	universal normal control
	PRO1054	lung tumor	universal normal control
	PRO1054	breast tumor	universal normal control
	PRO994	colon tumor	universal normal control
40	PRO994	lung tumor	universal normal control
	PRO994	rectal tumor	universal normal control
	PRO1069	lung tumor	universal normal control
	PRO1069	breast tumor	universal normal control
	PRO1411	colon tumor	universal normal control
45	PRO1411	lung tumor	universal normal control
	PRO1129	lung tumor	universal normal control
	PRO1129	rectal tumor	universal normal control
	PRO1359	colon tumor	universal normal control
	PRO1359	lung tumor	universal normal control
50	PRO1359	breast tumor	universal normal control
	PRO1359	prostate tumor	universal normal control
	PRO1139	lung tumor	universal normal control
	PRO1065	lung tumor	universal normal control
	PRO1028	colon tumor	universal normal control
55	PRO1028	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1028	breast tumor	universal normal control
	PRO1028	cervical tumor	universal normal control
	PRO1027	colon tumor	universal normal control
5	PRO1027	lung tumor	universal normal control
	PRO1027	breast tumor	universal normal control
	PRO1140	colon tumor	universal normal control
	PRO1140	breast tumor	universal normal control
	PRO1291	colon tumor	universal normal control
10	PRO1291	breast tumor	universal normal control
	PRO1105	colon tumor	universal normal control
	PRO1105	lung tumor	universal normal control
	PRO1026	lung tumor	universal normal control
	PRO1026	prostate tumor	universal normal control
15	PRO1104	colon tumor	universal normal control
	PRO1104	lung tumor	universal normal control
	PRO1104	breast tumor	universal normal control
	PRO1100	colon tumor	universal normal control
	PRO1100	lung tumor	universal normal control
20	PRO1100	breast tumor	universal normal control
	PRO1100	rectal tumor	universal normal control
	PRO1141	lung tumor	universal normal control
	PRO1772	colon tumor	universal normal control
	PRO1772	lung tumor	universal normal control
25	PRO1772	breast tumor	universal normal control
	PRO1772	cervical tumor	universal normal control
	PRO1064	colon tumor	universal normal control
	PRO1064	lung tumor	universal normal control
	PRO1379	colon tumor	universal normal control
30	PRO1379	lung tumor	universal normal control
	PRO1379	cervical tumor	universal normal control
	PRO3573	lung tumor	universal normal control
	PRO3573	breast tumor	universal normal control
	PRO3566	colon tumor	universal normal control
35	PRO3566	lung tumor	universal normal control
	PRO1156	lung tumor	universal normal control
	PRO1156	breast tumor	universal normal control
	PRO1156	prostate tumor	universal normal control
	PRO1098	colon tumor	universal normal control
40	PRO1098	lung tumor	universal normal control
	PRO1098	rectal tumor	universal normal control
	PRO1128	colon tumor	universal normal control
	PRO1128	lung tumor	universal normal control
	PRO1128	breast tumor	universal normal control
45	PRO1248	lung tumor	universal normal control
	PRO1248	breast tumor	universal normal control
	PRO1127	colon tumor	universal normal control
	PRO1127	lung tumor	universal normal control
	PRO1127	breast tumor	universal normal control
50	PRO1316	colon tumor	universal normal control
	PRO1316	lung tumor	universal normal control
	PRO1316	breast tumor	universal normal control
	PRO1197	colon tumor	universal normal control
	PRO1197	lung tumor	universal normal control
55	PRO1197	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1125	lung tumor	universal normal control
	PRO1158	breast tumor	universal normal control
	PRO1124	colon tumor	universal normal control
5	PRO1124	lung tumor	universal normal control
	PRO1380	colon tumor	universal normal control
	PRO1380	lung tumor	universal normal control
	PRO1380	breast tumor	universal normal control
	PRO1380	liver tumor	universal normal control
10	PRO1377	colon tumor	universal normal control
	PRO1377	lung tumor	universal normal control
	PRO1287	lung tumor	universal normal control
	PRO1287	breast tumor	universal normal control
	PRO1249	lung tumor	universal normal control
15	PRO1249	breast tumor	universal normal control
	PRO1335	colon tumor	universal normal control
	PRO1335	lung tumor	universal normal control
	PRO1335	breast tumor	universal normal control
	PRO3572	lung tumor	universal normal control
20	PRO1599	colon tumor	universal normal control
	PRO1599	lung tumor	universal normal control
	PRO1599	breast tumor	universal normal control
	PRO1374	lung tumor	universal normal control
	PRO1374	breast tumor	universal normal control
25	PRO1345	lung tumor	universal normal control
	PRO1345	breast tumor	universal normal control
	PRO1311	lung tumor	universal normal control
	PRO1311	breast tumor	universal normal control
	PRO1357	colon tumor	universal normal control
30	PRO1357	lung tumor	universal normal control
	PRO1557	colon tumor	universal normal control
	PRO1557	lung tumor	universal normal control
	PRO1557	breast tumor	universal normal control
	PRO1305	colon tumor	universal normal control
35	PRO1305	lung tumor	universal normal control
	PRO1305	breast tumor	universal normal control
	PRO1302	colon tumor	universal normal control
	PRO1302	lung tumor	universal normal control
	PRO1302	breast tumor	universal normal control
40	PRO1302	rectal tumor	universal normal control
	PRO1266	colon tumor	universal normal control
	PRO1336	colon tumor	universal normal control
	PRO1336	lung tumor	universal normal control
	PRO1336	breast tumor	universal normal control
45	PRO1278	colon tumor	universal normal control
	PRO1278	lung tumor	universal normal control
	PRO1270	breast tumor	universal normal control
	PRO1298	colon tumor	universal normal control
	PRO1298	lung tumor	universal normal control
50	PRO1301	lung tumor	universal normal control
	PRO1301	breast tumor	universal normal control
	PRO1268	colon tumor	universal normal control
	PRO1268	breast tumor	universal normal control
	PRO1327	lung tumor	universal normal control
55	PRO1327	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1328	colon tumor	universal normal control
	PRO1328	lung tumor	universal normal control
	PRO1328	breast tumor	universal normal control
5	PRO1329	colon tumor	universal normal control
	PRO1329	lung tumor	universal normal control
	PRO1329	breast tumor	universal normal control
	PRO1339	colon tumor	universal normal control
	PRO1339	lung tumor	universal normal control
10	PRO1342	colon tumor	universal normal control
	PRO1342	lung tumor	universal normal control
	PRO1342	breast tumor	universal normal control
	PRO1342	rectal tumor	universal normal control
	PRO1487	colon tumor	universal normal control
15	PRO1487	breast tumor	universal normal control
	PRO3579	lung tumor	universal normal control
	PRO3579	breast tumor	universal normal control
	PRO1472	colon tumor	universal normal control
	PRO1472	lung tumor	universal normal control
20	PRO1385	lung tumor	universal normal control
	PRO1385	breast tumor	universal normal control
	PRO1461	colon tumor	universal normal control
	PRO1461	lung tumor	universal normal control
	PRO1461	breast tumor	universal normal control
25	PRO1429	colon tumor	universal normal control
	PRO1429	lung tumor	universal normal control
	PRO1429	breast tumor	universal normal control
	PRO1568	lung tumor	universal normal control
	PRO1568	breast tumor	universal normal control
30	PRO1569	colon tumor	universal normal control
	PRO1569	lung tumor	universal normal control
	PRO1569	breast tumor	universal normal control
	PRO1753	colon tumor	universal normal control
	PRO1753	lung tumor	universal normal control
35	PRO1570	colon tumor	universal normal control
	PRO1570	lung tumor	universal normal control
	PRO1570	breast tumor	universal normal control
	PRO1570	prostate tumor	universal normal control
	PRO1570	rectal tumor	universal normal control
40	PRO1559	colon tumor	universal normal control
	PRO1559	lung tumor	universal normal control
	PRO1559	breast tumor	universal normal control
	PRO1486	lung tumor	universal normal control
	PRO1486	breast tumor	universal normal control
45	PRO1433	colon tumor	universal normal control
	PRO1433	lung tumor	universal normal control
	PRO1433	breast tumor	universal normal control
	PRO1433	rectal tumor	universal normal control
	PRO1490	lung tumor	universal normal control
50	PRO1490	breast tumor	universal normal control
	PRO1482	lung tumor	universal normal control
	PRO1482	breast tumor	universal normal control
	PRO1409	colon tumor	universal normal control
	PRO1409	lung tumor	universal normal control
55	PRO1409	breast tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1446	colon tumor	universal normal control
	PRO1446	lung tumor	universal normal control
	PRO1446	breast tumor	universal normal control
5	PRO1446	prostate tumor	universal normal control
	PRO1604	colon tumor	universal normal control
	PRO1604	lung tumor	universal normal control
	PRO1604	breast tumor	universal normal control
	PRO1491	colon tumor	universal normal control
10	PRO1491	lung tumor	universal normal control
	PRO1491	breast tumor	universal normal control
	PRO1431	colon tumor	universal normal control
	PRO1431	lung tumor	universal normal control
	PRO1563	colon tumor	universal normal control
15	PRO1563	lung tumor	universal normal control
	PRO1563	breast tumor	universal normal control
	PRO1571	colon tumor	universal normal control
	PRO1571	lung tumor	universal normal control
	PRO1571	breast tumor	universal normal control
20	PRO1572	lung tumor	universal normal control
	PRO1572	prostate tumor	universal normal control
	PRO1573	lung tumor	universal normal control
	PRO1573	breast tumor	universal normal control
	PRO1508	lung tumor	universal normal control
25	PRO1508	breast tumor	universal normal control
	PRO1485	colon tumor	universal normal control
	PRO1485	lung tumor	universal normal control
	PRO1564	colon tumor	universal normal control
	PRO1564	lung tumor	universal normal control
30	PRO1564	breast tumor	universal normal control
	PRO1550	colon tumor	universal normal control
	PRO1550	lung tumor	universal normal control
	PRO1550	breast tumor	universal normal control
	PRO1757	lung tumor	universal normal control
35	PRO1757	breast tumor	universal normal control
	PRO1757	prostate tumor	universal normal control
	PRO1758	lung tumor	universal normal control
	PRO1781	colon tumor	universal normal control
	PRO1781	lung tumor	universal normal control
40	PRO1781	breast tumor	universal normal control
	PRO1606	lung tumor	universal normal control
	PRO1606	breast tumor	universal normal control
	PRO1784	colon tumor	universal normal control
	PRO1784	lung tumor	universal normal control
45	PRO1784	breast tumor	universal normal control
	PRO1774	colon tumor	universal normal control
	PRO1774	lung tumor	universal normal control
	PRO1774	breast tumor	universal normal control
	PRO1605	colon tumor	universal normal control
50	PRO1605	lung tumor	universal normal control
	PRO1605	prostate tumor	universal normal control
	PRO1928	colon tumor	universal normal control
	PRO1928	lung tumor	universal normal control
	PRO1928	cervical tumor	universal normal control
55	PRO1865	lung tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO1865	liver tumor	universal normal control
	PRO1925	lung tumor	universal normal control
	PRO1926	liver tumor	universal normal control
5	PRO2630	colon tumor	universal normal control
	PRO2630	lung tumor	universal normal control
	PRO2630	breast tumor	universal normal control
	PRO2630	liver tumor	universal normal control
	PRO3443	colon tumor	universal normal control
10	PRO3443	lung tumor	universal normal control
	PRO3443	breast tumor	universal normal control
	PRO3301	colon tumor	universal normal control
	PRO3301	lung tumor	universal normal control
	PRO3301	breast tumor	universal normal control
15	PRO3301	rectal tumor	universal normal control
	PRO3442	colon tumor	universal normal control
	PRO3442	lung tumor	universal normal control
	PRO3442	rectal tumor	universal normal control
	PRO4978	colon tumor	universal normal control
20	PRO4978	lung tumor	universal normal control
	PRO4978	breast tumor	universal normal control
	PRO4978	rectal tumor	universal normal control
	PRO5801	colon tumor	universal normal control
	PRO5801	breast tumor	universal normal control
25	PRO19630	colon tumor	universal normal control
	PRO203	colon tumor	universal normal control
	PRO204	colon tumor	universal normal control
	PRO204	lung tumor	universal normal control
	PRO204	breast tumor	universal normal control
30	PRO204	prostate tumor	universal normal control
	PRO210	colon tumor	universal normal control
	PRO210	lung tumor	universal normal control
	PRO223	lung tumor	universal normal control
	PRO223	breast tumor	universal normal control
35	PRO247	colon tumor	universal normal control
	PRO247	lung tumor	universal normal control
	PRO247	breast	universal normal control
	PRO358	lung tumor	universal normal control
	PRO358	breast tumor	universal normal control
40	PRO358	prostate tumor	universal normal control
	PRO724	lung tumor	universal normal control
	PRO868	colon tumor	universal normal control
	PRO868	lung tumor	universal normal control
	PRO868	prostate tumor	universal normal control
45	PRO868	rectal tumor	universal normal control
	PRO740	colon tumor	universal normal control
	PRO1478	colon tumor	universal normal control
	PRO1478	lung tumor	universal normal control
	PRO162	colon tumor	universal normal control
50	PRO162	lung tumor	universal normal control
	PRO162	breast tumor	universal normal control
	PRO828	colon tumor	universal normal control
	PRO828	lung tumor	universal normal control
	PRO828	breast tumor	universal normal control
55	PRO828	cervical tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO828	liver tumor	universal normal control
	PRO819	lung tumor	universal normal control
	PRO819	breast tumor	universal normal control
5	PRO819	rectal tumor	universal normal control
	PRO813	colon tumor	universal normal control
	PRO813	lung tumor	universal normal control
	PRO813	breast tumor	universal normal control
	PRO813	prostate tumor	universal normal control
10	PRO1194	colon tumor	universal normal control
	PRO1194	lung tumor	universal normal control
	PRO1194	breast tumor	universal normal control
	PRO887	colon tumor	universal normal control
	PRO887	lung tumor	universal normal control
15	PRO887	rectal tumor	universal normal control
	PRO1071	colon tumor	universal normal control
	PRO1071	lung tumor	universal normal control
	PRO1071	breast tumor	universal normal control
	PRO1029	colon tumor	universal normal control
20	PRO1029	lung tumor	universal normal control
	PRO1029	breast tumor	universal normal control
	PRO1190	lung tumor	universal normal control
	PRO1190	breast tumor	universal normal control
	PRO4334	lung tumor	universal normal control
25	PRO1155	colon tumor	universal normal control
	PRO1155	lung tumor	universal normal control
	PRO1157	breast tumor	universal normal control
	PRO1157	cervical tumor	universal normal control
	PRO1122	lung tumor	universal normal control
30	PRO1122	breast tumor	universal normal control
	PRO1183	colon tumor	universal normal control
	PRO1183	lung tumor	universal normal control
	PRO1183	breast tumor	universal normal control
	PRO1337	colon tumor	universal normal control
35	PRO1337	lung tumor	universal normal control
	PRO1337	breast tumor	universal normal control
	PRO1480	colon tumor	universal normal control
	PRO1480	lung tumor	universal normal control
	PRO1480	breast tumor	universal normal control
40	PRO19645	colon tumor	universal normal control
	PRO9782	colon tumor	universal normal control
	PRO1419	colon tumor	universal normal control
	PRO1575	colon tumor	universal normal control
	PRO1575	lung tumor	universal normal control
45	PRO1567	colon tumor	universal normal control
	PRO1567	lung tumor	universal normal control
	PRO1567	breast tumor	universal normal control
	PRO1891	colon tumor	universal normal control
	PRO1889	colon tumor	universal normal control
50	PRO1889	lung tumor	universal normal control
	PRO1785	lung tumor	universal normal control
	PRO1785	prostate tumor	universal normal control
	PRO6003	colon tumor	universal normal control
	PRO4333	colon tumor	universal normal control
55	PRO4356	colon tumor	universal normal control



Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO4352	colon tumor	universal normal control
	PRO4354	colon tumor	universal normal control
	PRO4354	lung tumor	universal normal control
5	PRO4354	prostate tumor	universal normal control
	PRO4369	colon tumor	universal normal control
	PRO6030	colon tumor	universal normal control
	PRO4433	colon tumor	universal normal control
	PRO4424	colon tumor	universal normal control
10	PRO4424	breast tumor	universal normal control
	PRO6017	colon tumor	universal normal control
	PRO19563	colon tumor	universal normal control
	PRO6015	colon tumor	universal normal control
	PRO5779	colon tumor	universal normal control
15	PRO5776	colon tumor	universal normal control
	PRO4430	lung tumor	universal normal control
	PRO4421	colon tumor	universal normal control
	PRO4499	colon tumor	universal normal control
	PRO4423	colon tumor	universal normal control
20	PRO5998	colon tumor	universal normal control
	PRO5998	lung tumor	universal normal control
	PRO4501	colon tumor	universal normal control
	PRO6240	colon tumor	universal normal control
	PRO6245	colon tumor	universal normal control
25	PRO6175	colon tumor	universal normal control
	PRO9742	colon tumor	universal normal control
	PRO7179	colon tumor	universal normal control
	PRO6239	colon tumor	universal normal control
	PRO6493	colon tumor	universal normal control
30	PRO9741	colon tumor	universal normal control
	PRO9822	colon tumor	universal normal control
	PRO6244	colon tumor	universal normal control
	PRO9740	colon tumor	universal normal control
	PRO9739	colon tumor	universal normal control
35	PRO7177	colon tumor	universal normal control
	PRO7178	colon tumor	universal normal control
	PRO6246	colon tumor	universal normal control
	PRO6241	colon tumor	universal normal control
	PRO9835	colon tumor	universal normal control
40	PRO9857	colon tumor	universal normal control
	PRO7436	colon tumor	universal normal control
	PRO9856	colon tumor	universal normal control
	PRO19605	colon tumor	universal normal control
	PRO9859	colon tumor	universal normal control
45	PRO12970	colon tumor	universal normal control
	PRO19626	colon tumor	universal normal control
	PRO9883	colon tumor	universal normal control
	PRO19670	colon tumor	universal normal control
	PRO19624	colon tumor	universal normal control
50	PRO19680	colon tumor	universal normal control
	PRO19675	colon tumor	universal normal control
	PRO9834	colon tumor	universal normal control
	PRO9744	colon tumor	universal normal control
	PRO19644	colon tumor	universal normal control
55	PRO19625	colon tumor	universal normal control

Table 8 (cont')

	<u>Molecule</u>	<u>is overexpressed in:</u>	<u>as compared to:</u>
	PRO19597	colon tumor	universal normal control
	PRO16090	colon tumor	universal normal control
	PRO19576	colon tumor	universal normal control
5	PRO19646	colon tumor	universal normal control
	PRO19814	colon tumor	universal normal control
	PRO19669	colon tumor	universal normal control
	PRO19818	colon tumor	universal normal control
	PRO20088	colon tumor	universal normal control
10	PRO16089	colon tumor	universal normal control
	PRO20025	colon tumor	universal normal control
	PRO20040	colon tumor	universal normal control
	PRO1760	adrenal tumor	universal normal control
	PRO1760	breast tumor	universal normal control
15	PRO1760	cervical tumor	universal normal control
	PRO1760	colon tumor	universal normal control
	PRO1760	liver tumor	universal normal control
	PRO1760	lung tumor	universal normal control
	PRO1760	prostate tumor	universal normal control
20	PRO1760	rectal tumor	universal normal control
	PRO6029	adrenal tumor	universal normal control
	PRO6029	colon tumor	universal normal control
	PRO6029	prostate tumor	universal normal control
	PRO1801	colon tumor	universal normal control
25	PRO1801	lung tumor	universal normal control



[illegible]

NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), and Figure 610 (SEQ ID NO:610).

2. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure 71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figures 75A-75B (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID

NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure 221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225), Figure 227 (SEQ ID NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ ID NO:233), Figure 235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239), Figure 241 (SEQ ID NO:241), Figure 243 (SEQ ID NO:243), Figure 245 (SEQ ID NO:245), Figure 247 (SEQ ID NO:247), Figure 249 (SEQ ID NO:249), Figure 251 (SEQ ID NO:251), Figure 253 (SEQ ID NO:253), Figure 255 (SEQ ID NO:255), Figure 257 (SEQ ID NO:257), Figure 259 (SEQ ID NO:259), Figure 261 (SEQ ID NO:261), Figure 263 (SEQ ID NO:263), Figure 265 (SEQ ID NO:265), Figure 267 (SEQ ID NO:267), Figure 269 (SEQ ID NO:269), Figure 271 (SEQ ID NO:271), Figure 273 (SEQ ID NO:273), Figure 275 (SEQ ID NO:275), Figure 277 (SEQ ID NO:277), Figure 279 (SEQ ID NO:279), Figure 281 (SEQ ID NO:281), Figure 283 (SEQ ID NO:283), Figure 285 (SEQ ID NO:285), Figure 287 (SEQ ID NO:287), Figures 289A-289B (SEQ ID NO:289), Figure 291 (SEQ ID NO:291), Figure 293 (SEQ ID NO:293), Figure 295 (SEQ ID NO:295), Figure 297 (SEQ ID NO:297), Figure 299 (SEQ ID NO:299), Figure 301 (SEQ ID NO:301), Figure 303 (SEQ ID NO:303), Figure 305 (SEQ ID NO:305), Figure 307 (SEQ ID NO:307), Figure 309 (SEQ ID NO:309), Figures 311A-311B (SEQ ID NO:311), Figure 313 (SEQ ID NO:313), Figure 315 (SEQ ID NO:315), Figure 317 (SEQ ID NO:317), Figure 319 (SEQ ID NO:319), Figure 321 (SEQ ID NO:321), Figure 323 (SEQ ID NO:323), Figure 325 (SEQ ID NO:325), Figure 327 (SEQ ID NO:327), Figure 329 (SEQ ID NO:329), Figure 331 (SEQ ID NO:331), Figure 333 (SEQ ID NO:333), Figure 335 (SEQ ID NO:335), Figure 337 (SEQ ID NO:337), Figure 339 (SEQ ID NO:339), Figure 341 (SEQ ID NO:341), Figure 343 (SEQ ID NO:343), Figure 345 (SEQ ID NO:345), Figure 347 (SEQ ID NO:347), Figure 349 (SEQ ID NO:349), Figures 351A-351B (SEQ ID NO:351), Figure 353 (SEQ ID NO:353), Figure 355 (SEQ ID NO:355), Figure 357 (SEQ ID NO:357), Figure 359 (SEQ ID NO:359), Figure 361 (SEQ ID NO:361), Figure 363 (SEQ ID NO:363), Figure 365 (SEQ ID NO:365), Figure 367 (SEQ ID NO:367), Figure 369 (SEQ ID NO:369), Figure 371 (SEQ ID NO:371), Figure 373 (SEQ ID NO:373), Figure 375 (SEQ ID NO:375), Figure 377 (SEQ ID NO:377), Figure 379 (SEQ ID NO:379), Figure 381 (SEQ ID NO:381), Figure 383 (SEQ ID NO:383), Figure 385 (SEQ ID NO:385), Figure 387 (SEQ ID NO:387), Figure 389 (SEQ ID NO:389), Figure 391 (SEQ ID NO:391), Figure 393 (SEQ ID NO:393), Figure 395 (SEQ ID NO:395), Figure 397 (SEQ ID NO:397), Figure 399 (SEQ ID NO:399), Figure 401 (SEQ ID NO:401), Figure 403 (SEQ ID NO:403), Figure 405 (SEQ ID NO:405), Figure 407 (SEQ ID NO:407), Figure 409 (SEQ ID NO:409), Figure 411 (SEQ ID NO:411), Figure 413 (SEQ ID NO:413), Figure 415 (SEQ ID NO:415), Figure 417 (SEQ ID NO:417), Figure 419 (SEQ ID NO:419), Figure 421 (SEQ ID NO:421), Figure 423 (SEQ ID NO:423), Figure 425 (SEQ ID NO:425), Figure 427 (SEQ ID NO:427), Figure 429 (SEQ ID NO:429), Figure 431 (SEQ ID

NO:431), Figure 433 (SEQ ID NO:433), Figure 435 (SEQ ID NO:435), Figure 437 (SEQ ID NO:437), Figure 439 (SEQ ID NO:439), Figure 441 (SEQ ID NO:441), Figure 443 (SEQ ID NO:443), Figure 445 (SEQ ID NO:445), Figure 447 (SEQ ID NO:447), Figure 449 (SEQ ID NO:449), Figure 451 (SEQ ID NO:451), Figure 453 (SEQ ID NO:453), Figure 455 (SEQ ID NO:455), Figure 457 (SEQ ID NO:457), Figure 459 (SEQ ID NO:459), Figure 461 (SEQ ID NO:461), Figure 463 (SEQ ID NO:463), Figure 465 (SEQ ID NO:465), Figure 467 (SEQ ID NO:467), Figure 469 (SEQ ID NO:469), Figure 471 (SEQ ID NO:471), Figure 473 (SEQ ID NO:473), Figure 475 (SEQ ID NO:475), Figure 477 (SEQ ID NO:477), Figure 479 (SEQ ID NO:479), Figure 481 (SEQ ID NO:481), Figure 483 (SEQ ID NO:483), Figure 485 (SEQ ID NO:485), Figure 487 (SEQ ID NO:487), Figure 489 (SEQ ID NO:489), Figure 491 (SEQ ID NO:491), Figure 493 (SEQ ID NO:493), Figure 495 (SEQ ID NO:495), Figure 497 (SEQ ID NO:497), Figure 499 (SEQ ID NO:499), Figure 501 (SEQ ID NO:501), Figure 503 (SEQ ID NO:503), Figure 505 (SEQ ID NO:505), Figure 507 (SEQ ID NO:507), Figure 509 (SEQ ID NO:509), Figure 511 (SEQ ID NO:511), Figure 513 (SEQ ID NO:513), Figure 515 (SEQ ID NO:515), Figure 517 (SEQ ID NO:517), Figure 519 (SEQ ID NO:519), Figure 521 (SEQ ID NO:521), Figure 523 (SEQ ID NO:523), Figures 525A-525B (SEQ ID NO:525), Figure 527 (SEQ ID NO:527), Figure 529 (SEQ ID NO:529), Figure 531 (SEQ ID NO:531), Figure 533 (SEQ ID NO:533), Figure 535 (SEQ ID NO:535), Figure 537 (SEQ ID NO:537), Figure 539 (SEQ ID NO:539), Figure 541 (SEQ ID NO:541), Figure 543 (SEQ ID NO:543), Figure 545 (SEQ ID NO:545), Figure 547 (SEQ ID NO:547), Figure 549 (SEQ ID NO:549), Figure 551 (SEQ ID NO:551), Figure 553 (SEQ ID NO:553), Figure 555 (SEQ ID NO:555), Figure 557 (SEQ ID NO:557), Figure 559 (SEQ ID NO:559), Figure 561 (SEQ ID NO:561), Figure 563 (SEQ ID NO:563), Figure 565 (SEQ ID NO:565), Figure 567 (SEQ ID NO:567), Figure 569 (SEQ ID NO:569), Figure 571 (SEQ ID NO:571), Figure 573 (SEQ ID NO:573), Figure 575 (SEQ ID NO:575), Figure 577 (SEQ ID NO:577), Figure 579 (SEQ ID NO:579), Figure 581 (SEQ ID NO:581), Figure 583 (SEQ ID NO:583), Figure 585 (SEQ ID NO:585), Figure 587 (SEQ ID NO:587), Figure 589 (SEQ ID NO:589), Figure 591 (SEQ ID NO:591), Figure 593 (SEQ ID NO:593), Figure 595 (SEQ ID NO:595), Figure 597 (SEQ ID NO:597), Figure 599 (SEQ ID NO:599), Figure 601 (SEQ ID NO:601), Figure 603 (SEQ ID NO:603), Figure 605 (SEQ ID NO:605), Figure 607 (SEQ ID NO:607), and Figure 609 (SEQ ID NO:609).

3. Isolated nucleic acid having at least 80% nucleic acid sequence identity to a nucleotide sequence selected from the group consisting of the full-length coding sequence of the nucleotide sequence shown in Figure 1 (SEQ ID NO:1), Figure 3 (SEQ ID NO:3), Figure 5 (SEQ ID NO:5), Figure 7 (SEQ ID NO:7), Figure 9 (SEQ ID NO:9), Figure 11 (SEQ ID NO:11), Figure 13 (SEQ ID NO:13), Figure 15 (SEQ ID NO:15), Figure 17 (SEQ ID NO:17), Figure 19 (SEQ ID NO:19), Figure 21 (SEQ ID NO:21), Figure 23 (SEQ ID NO:23), Figure 25 (SEQ ID NO:25), Figure 27 (SEQ ID NO:27), Figure 29 (SEQ ID NO:29), Figure 31 (SEQ ID NO:31), Figure 33 (SEQ ID NO:33), Figure 35 (SEQ ID NO:35), Figure 37 (SEQ ID NO:37), Figure 39 (SEQ ID NO:39), Figure 41 (SEQ ID NO:41), Figure 43 (SEQ ID NO:43), Figure 45 (SEQ ID NO:45), Figure 47 (SEQ ID NO:47), Figure 49 (SEQ ID NO:49), Figure 51 (SEQ ID NO:51), Figure 53 (SEQ ID NO:53), Figure 55 (SEQ ID NO:55), Figure 57 (SEQ ID NO:57), Figure 59 (SEQ ID NO:59), Figure 61 (SEQ ID NO:61), Figure 63 (SEQ ID NO:63), Figure 65 (SEQ ID NO:65), Figure 67 (SEQ ID NO:67), Figure 69 (SEQ ID NO:69), Figure

71 (SEQ ID NO:71), Figure 73 (SEQ ID NO:73), Figures 75A-75B (SEQ ID NO:75), Figure 77 (SEQ ID NO:77), Figure 79 (SEQ ID NO:79), Figure 81 (SEQ ID NO:81), Figure 83 (SEQ ID NO:83), Figure 85 (SEQ ID NO:85), Figure 87 (SEQ ID NO:87), Figure 89 (SEQ ID NO:89), Figure 91 (SEQ ID NO:91), Figure 93 (SEQ ID NO:93), Figure 95 (SEQ ID NO:95), Figure 97 (SEQ ID NO:97), Figure 99 (SEQ ID NO:99), Figure 101 (SEQ ID NO:101), Figure 103 (SEQ ID NO:103), Figure 105 (SEQ ID NO:105), Figure 107 (SEQ ID NO:107), Figure 109 (SEQ ID NO:109), Figure 111 (SEQ ID NO:111), Figure 113 (SEQ ID NO:113), Figure 115 (SEQ ID NO:115), Figure 117 (SEQ ID NO:117), Figure 119 (SEQ ID NO:119), Figure 121 (SEQ ID NO:121), Figure 123 (SEQ ID NO:123), Figure 125 (SEQ ID NO:125), Figure 127 (SEQ ID NO:127), Figure 129 (SEQ ID NO:129), Figure 131 (SEQ ID NO:131), Figure 133 (SEQ ID NO:133), Figure 135 (SEQ ID NO:135), Figure 137 (SEQ ID NO:137), Figure 139 (SEQ ID NO:139), Figure 141 (SEQ ID NO:141), Figure 143 (SEQ ID NO:143), Figure 145 (SEQ ID NO:145), Figure 147 (SEQ ID NO:147), Figure 149 (SEQ ID NO:149), Figure 151 (SEQ ID NO:151), Figure 153 (SEQ ID NO:153), Figure 155 (SEQ ID NO:155), Figure 157 (SEQ ID NO:157), Figure 159 (SEQ ID NO:159), Figure 161 (SEQ ID NO:161), Figure 163 (SEQ ID NO:163), Figure 165 (SEQ ID NO:165), Figure 167 (SEQ ID NO:167), Figure 169 (SEQ ID NO:169), Figure 171 (SEQ ID NO:171), Figure 173 (SEQ ID NO:173), Figure 175 (SEQ ID NO:175), Figure 177 (SEQ ID NO:177), Figure 179 (SEQ ID NO:179), Figure 181 (SEQ ID NO:181), Figure 183 (SEQ ID NO:183), Figure 185 (SEQ ID NO:185), Figure 187 (SEQ ID NO:187), Figure 189 (SEQ ID NO:189), Figure 191 (SEQ ID NO:191), Figure 193 (SEQ ID NO:193), Figure 195 (SEQ ID NO:195), Figure 197 (SEQ ID NO:197), Figure 199 (SEQ ID NO:199), Figure 201 (SEQ ID NO:201), Figure 203 (SEQ ID NO:203), Figure 205 (SEQ ID NO:205), Figure 207 (SEQ ID NO:207), Figure 209 (SEQ ID NO:209), Figure 211 (SEQ ID NO:211), Figure 213 (SEQ ID NO:213), Figure 215 (SEQ ID NO:215), Figure 217 (SEQ ID NO:217), Figure 219 (SEQ ID NO:219), Figure 221 (SEQ ID NO:221), Figure 223 (SEQ ID NO:223), Figure 225 (SEQ ID NO:225), Figure 227 (SEQ ID NO:227), Figure 229 (SEQ ID NO:229), Figure 231 (SEQ ID NO:231), Figure 233 (SEQ ID NO:233), Figure 235 (SEQ ID NO:235), Figure 237 (SEQ ID NO:237), Figure 239 (SEQ ID NO:239), Figure 241 (SEQ ID NO:241), Figure 243 (SEQ ID NO:243), Figure 245 (SEQ ID NO:245), Figure 247 (SEQ ID NO:247), Figure 249 (SEQ ID NO:249), Figure 251 (SEQ ID NO:251), Figure 253 (SEQ ID NO:253), Figure 255 (SEQ ID NO:255), Figure 257 (SEQ ID NO:257), Figure 259 (SEQ ID NO:259), Figure 261 (SEQ ID NO:261), Figure 263 (SEQ ID NO:263), Figure 265 (SEQ ID NO:265), Figure 267 (SEQ ID NO:267), Figure 269 (SEQ ID NO:269), Figure 271 (SEQ ID NO:271), Figure 273 (SEQ ID NO:273), Figure 275 (SEQ ID NO:275), Figure 277 (SEQ ID NO:277), Figure 279 (SEQ ID NO:279), Figure 281 (SEQ ID NO:281), Figure 283 (SEQ ID NO:283), Figure 285 (SEQ ID NO:285), Figure 287 (SEQ ID NO:287), Figures 289A-289B (SEQ ID NO:289), Figure 291 (SEQ ID NO:291), Figure 293 (SEQ ID NO:293), Figure 295 (SEQ ID NO:295), Figure 297 (SEQ ID NO:297), Figure 299 (SEQ ID NO:299), Figure 301 (SEQ ID NO:301), Figure 303 (SEQ ID NO:303), Figure 305 (SEQ ID NO:305), Figure 307 (SEQ ID NO:307), Figure 309 (SEQ ID NO:309), Figures 311A-311B (SEQ ID NO:311), Figure 313 (SEQ ID NO:313), Figure 315 (SEQ ID NO:315), Figure 317 (SEQ ID NO:317), Figure 319 (SEQ ID NO:319), Figure 321 (SEQ ID NO:321), Figure 323 (SEQ ID NO:323), Figure 325 (SEQ ID NO:325), Figure 327 (SEQ ID NO:327), Figure 329 (SEQ ID NO:329), Figure 331 (SEQ ID NO:331), Figure 333 (SEQ ID NO:333), Figure 335 (SEQ ID NO:335), Figure 337 (SEQ ID NO:337), Figure



339 (SEQ ID NO:339), Figure 341 (SEQ ID NO:341), Figure 343 (SEQ ID NO:343), Figure 345 (SEQ ID NO:345), Figure 347 (SEQ ID NO:347), Figure 349 (SEQ ID NO:349), Figures 351A-351B (SEQ ID NO:351), Figure 353 (SEQ ID NO:353), Figure 355 (SEQ ID NO:355), Figure 357 (SEQ ID NO:357), Figure 359 (SEQ ID NO:359), Figure 361 (SEQ ID NO:361), Figure 363 (SEQ ID NO:363), Figure 365 (SEQ ID NO:365), Figure 367 (SEQ ID NO:367), Figure 369 (SEQ ID NO:369), Figure 371 (SEQ ID NO:371), Figure 373 (SEQ ID NO:373), Figure 375 (SEQ ID NO:375), Figure 377 (SEQ ID NO:377), Figure 379 (SEQ ID NO:379), Figure 381 (SEQ ID NO:381), Figure 383 (SEQ ID NO:383), Figure 385 (SEQ ID NO:385), Figure 387 (SEQ ID NO:387), Figure 389 (SEQ ID NO:389), Figure 391 (SEQ ID NO:391), Figure 393 (SEQ ID NO:393), Figure 395 (SEQ ID NO:395), Figure 397 (SEQ ID NO:397), Figure 399 (SEQ ID NO:399), Figure 401 (SEQ ID NO:401), Figure 403 (SEQ ID NO:403), Figure 405 (SEQ ID NO:405), Figure 407 (SEQ ID NO:407), Figure 409 (SEQ ID NO:409), Figure 411 (SEQ ID NO:411), Figure 413 (SEQ ID NO:413), Figure 415 (SEQ ID NO:415), Figure 417 (SEQ ID NO:417), Figure 419 (SEQ ID NO:419), Figure 421 (SEQ ID NO:421), Figure 423 (SEQ ID NO:423), Figure 425 (SEQ ID NO:425), Figure 427 (SEQ ID NO:427), Figure 429 (SEQ ID NO:429), Figure 431 (SEQ ID NO:431), Figure 433 (SEQ ID NO:433), Figure 435 (SEQ ID NO:435), Figure 437 (SEQ ID NO:437), Figure 439 (SEQ ID NO:439), Figure 441 (SEQ ID NO:441), Figure 443 (SEQ ID NO:443), Figure 445 (SEQ ID NO:445), Figure 447 (SEQ ID NO:447), Figure 449 (SEQ ID NO:449), Figure 451 (SEQ ID NO:451), Figure 453 (SEQ ID NO:453), Figure 455 (SEQ ID NO:455), Figure 457 (SEQ ID NO:457), Figure 459 (SEQ ID NO:459), Figure 461 (SEQ ID NO:461), Figure 463 (SEQ ID NO:463), Figure 465 (SEQ ID NO:465), Figure 467 (SEQ ID NO:467), Figure 469 (SEQ ID NO:469), Figure 471 (SEQ ID NO:471), Figure 473 (SEQ ID NO:473), Figure 475 (SEQ ID NO:475), Figure 477 (SEQ ID NO:477), Figure 479 (SEQ ID NO:479), Figure 481 (SEQ ID NO:481), Figure 483 (SEQ ID NO:483), Figure 485 (SEQ ID NO:485), Figure 487 (SEQ ID NO:487), Figure 489 (SEQ ID NO:489), Figure 491 (SEQ ID NO:491), Figure 493 (SEQ ID NO:493), Figure 495 (SEQ ID NO:495), Figure 497 (SEQ ID NO:497), Figure 499 (SEQ ID NO:499), Figure 501 (SEQ ID NO:501), Figure 503 (SEQ ID NO:503), Figure 505 (SEQ ID NO:505), Figure 507 (SEQ ID NO:507), Figure 509 (SEQ ID NO:509), Figure 511 (SEQ ID NO:511), Figure 513 (SEQ ID NO:513), Figure 515 (SEQ ID NO:515), Figure 517 (SEQ ID NO:517), Figure 519 (SEQ ID NO:519), Figure 521 (SEQ ID NO:521), Figure 523 (SEQ ID NO:523), Figures 525A-525B (SEQ ID NO:525), Figure 527 (SEQ ID NO:527), Figure 529 (SEQ ID NO:529), Figure 531 (SEQ ID NO:531), Figure 533 (SEQ ID NO:533), Figure 535 (SEQ ID NO:535), Figure 537 (SEQ ID NO:537), Figure 539 (SEQ ID NO:539), Figure 541 (SEQ ID NO:541), Figure 543 (SEQ ID NO:543), Figure 545 (SEQ ID NO:545), Figure 547 (SEQ ID NO:547), Figure 549 (SEQ ID NO:549), Figure 551 (SEQ ID NO:551), Figure 553 (SEQ ID NO:553), Figure 555 (SEQ ID NO:555), Figure 557 (SEQ ID NO:557), Figure 559 (SEQ ID NO:559), Figure 561 (SEQ ID NO:561), Figure 563 (SEQ ID NO:563), Figure 565 (SEQ ID NO:565), Figure 567 (SEQ ID NO:567), Figure 569 (SEQ ID NO:569), Figure 571 (SEQ ID NO:571), Figure 573 (SEQ ID NO:573), Figure 575 (SEQ ID NO:575), Figure 577 (SEQ ID NO:577), Figure 579 (SEQ ID NO:579), Figure 581 (SEQ ID NO:581), Figure 583 (SEQ ID NO:583), Figure 585 (SEQ ID NO:585), Figure 587 (SEQ ID NO:587), Figure 589 (SEQ ID NO:589), Figure 591 (SEQ ID NO:591), Figure 593 (SEQ ID NO:593), Figure 595 (SEQ ID NO:595), Figure 597 (SEQ ID NO:597), Figure 599 (SEQ ID NO:599), Figure 601 (SEQ ID NO:601), Figure 603 (SEQ ID NO:603), Figure

605 (SEQ ID NO:605), Figure 607 (SEQ ID NO:607), and Figure 609 (SEQ ID NO:609).

4. Isolated nucleic acid having at least 80% nucleic acid sequence identity to the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

5. A vector comprising the nucleic acid of Claim 1.

6. A host cell comprising the vector of Claim 5.

7. The host cell of Claim 6, wherein said cell is a CHO cell.

8. The host cell of Claim 6, wherein said cell is an *E. coli*.

9. The host cell of Claim 6, wherein said cell is a yeast cell.

10. A process for producing a PRO polypeptide comprising culturing the host cell of Claim 6 under conditions suitable for expression of said PRO polypeptide and recovering said PRO polypeptide from the cell culture.

11. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence selected from the group consisting of the amino acid sequence shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID

[illegible]

NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), and Figure 610 (SEQ ID NO:610).

12. An isolated polypeptide having at least 80% amino acid sequence identity to an amino acid sequence encoded by the full-length coding sequence of the DNA deposited under any ATCC accession number shown in Table 7.

13. A chimeric molecule comprising a polypeptide according to Claim 11 fused to a heterologous amino acid sequence.

14. The chimeric molecule of Claim 13, wherein said heterologous amino acid sequence is an epitope tag sequence.

15. The chimeric molecule of Claim 13, wherein said heterologous amino acid sequence is a Fc region of an immunoglobulin.

5

16. An antibody which specifically binds to a polypeptide according to Claim 11.

17. The antibody of Claim 16, wherein said antibody is a monoclonal antibody, a humanized antibody or a single-chain antibody.

10

18. Isolated nucleic acid having at least 80% nucleic acid sequence identity to:

(a) a nucleotide sequence encoding the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure

35

188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure

454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), or Figure 610 (SEQ ID NO:610), lacking its associated signal peptide;

(b) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ

[illegible]



NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), or Figure 610 (SEQ ID NO:610), with its associated signal peptide; or

(c) a nucleotide sequence encoding an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18

(SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID

NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID

NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), or Figure 610 (SEQ ID NO:610), lacking its associated signal peptide.

10            19.        An isolated polypeptide having at least 80% amino acid sequence identity to:

(a)          an amino acid sequence of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ  
ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID  
NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ  
ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28  
15 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure  
36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42),  
Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID  
NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ  
ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66  
20 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure  
74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80),  
Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID  
NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ  
ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure  
25 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID  
NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ ID NO:116), Figure  
118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure 124 (SEQ ID  
NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID NO:130), Figure  
132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure 138 (SEQ ID  
NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID NO:144), Figure  
30 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure 152 (SEQ ID  
NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID NO:158), Figure  
160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure 166 (SEQ ID  
NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID NO:172), Figure  
35 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure 180 (SEQ ID  
NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID NO:186), Figure  
188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure 194 (SEQ ID

NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure 208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure 236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure 264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure 278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure 306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure 320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure 334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure 348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure 376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID

NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), or Figure 610 (SEQ ID NO:610), lacking its associated signal peptide;

(b) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure 34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40), Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64 (SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78), Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102 (SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108), Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ

ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure  
124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID  
NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure  
138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID  
NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure  
5 152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID  
NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure  
166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID  
NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure  
180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID  
10 NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure  
194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID  
NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure  
208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID  
NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure  
15 222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID  
NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure  
236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID  
NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure  
250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID  
20 NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure  
264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID  
NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure  
278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID  
NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure  
25 292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID  
NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure  
306 (SEQ ID NO:306), Figure 308 (SEQ ID NO:308), Figure 310 (SEQ ID NO:310), Figure 312 (SEQ ID  
NO:312), Figure 314 (SEQ ID NO:314), Figure 316 (SEQ ID NO:316), Figure 318 (SEQ ID NO:318), Figure  
320 (SEQ ID NO:320), Figure 322 (SEQ ID NO:322), Figure 324 (SEQ ID NO:324), Figure 326 (SEQ ID  
30 NO:326), Figure 328 (SEQ ID NO:328), Figure 330 (SEQ ID NO:330), Figure 332 (SEQ ID NO:332), Figure  
334 (SEQ ID NO:334), Figure 336 (SEQ ID NO:336), Figure 338 (SEQ ID NO:338), Figure 340 (SEQ ID  
NO:340), Figure 342 (SEQ ID NO:342), Figure 344 (SEQ ID NO:344), Figure 346 (SEQ ID NO:346), Figure  
348 (SEQ ID NO:348), Figure 350 (SEQ ID NO:350), Figure 352 (SEQ ID NO:352), Figure 354 (SEQ ID  
NO:354), Figure 356 (SEQ ID NO:356), Figure 358 (SEQ ID NO:358), Figure 360 (SEQ ID NO:360), Figure  
35 362 (SEQ ID NO:362), Figure 364 (SEQ ID NO:364), Figure 366 (SEQ ID NO:366), Figure 368 (SEQ ID  
NO:368), Figure 370 (SEQ ID NO:370), Figure 372 (SEQ ID NO:372), Figure 374 (SEQ ID NO:374), Figure  
376 (SEQ ID NO:376), Figure 378 (SEQ ID NO:378), Figure 380 (SEQ ID NO:380), Figure 382 (SEQ ID

NO:382), Figure 384 (SEQ ID NO:384), Figure 386 (SEQ ID NO:386), Figure 388 (SEQ ID NO:388), Figure 390 (SEQ ID NO:390), Figure 392 (SEQ ID NO:392), Figure 394 (SEQ ID NO:394), Figure 396 (SEQ ID NO:396), Figure 398 (SEQ ID NO:398), Figure 400 (SEQ ID NO:400), Figure 402 (SEQ ID NO:402), Figure 404 (SEQ ID NO:404), Figure 406 (SEQ ID NO:406), Figure 408 (SEQ ID NO:408), Figure 410 (SEQ ID NO:410), Figure 412 (SEQ ID NO:412), Figure 414 (SEQ ID NO:414), Figure 416 (SEQ ID NO:416), Figure 418 (SEQ ID NO:418), Figure 420 (SEQ ID NO:420), Figure 422 (SEQ ID NO:422), Figure 424 (SEQ ID NO:424), Figure 426 (SEQ ID NO:426), Figure 428 (SEQ ID NO:428), Figure 430 (SEQ ID NO:430), Figure 432 (SEQ ID NO:432), Figure 434 (SEQ ID NO:434), Figure 436 (SEQ ID NO:436), Figure 438 (SEQ ID NO:438), Figure 440 (SEQ ID NO:440), Figure 442 (SEQ ID NO:442), Figure 444 (SEQ ID NO:444), Figure 446 (SEQ ID NO:446), Figure 448 (SEQ ID NO:448), Figure 450 (SEQ ID NO:450), Figure 452 (SEQ ID NO:452), Figure 454 (SEQ ID NO:454), Figure 456 (SEQ ID NO:456), Figure 458 (SEQ ID NO:458), Figure 460 (SEQ ID NO:460), Figure 462 (SEQ ID NO:462), Figure 464 (SEQ ID NO:464), Figure 466 (SEQ ID NO:466), Figure 468 (SEQ ID NO:468), Figure 470 (SEQ ID NO:470), Figure 472 (SEQ ID NO:472), Figure 474 (SEQ ID NO:474), Figure 476 (SEQ ID NO:476), Figure 478 (SEQ ID NO:478), Figure 480 (SEQ ID NO:480), Figure 482 (SEQ ID NO:482), Figure 484 (SEQ ID NO:484), Figure 486 (SEQ ID NO:486), Figure 488 (SEQ ID NO:488), Figure 490 (SEQ ID NO:490), Figure 492 (SEQ ID NO:492), Figure 494 (SEQ ID NO:494), Figure 496 (SEQ ID NO:496), Figure 498 (SEQ ID NO:498), Figure 500 (SEQ ID NO:500), Figure 502 (SEQ ID NO:502), Figure 504 (SEQ ID NO:504), Figure 506 (SEQ ID NO:506), Figure 508 (SEQ ID NO:508), Figure 510 (SEQ ID NO:510), Figure 512 (SEQ ID NO:512), Figure 514 (SEQ ID NO:514), Figure 516 (SEQ ID NO:516), Figure 518 (SEQ ID NO:518), Figure 520 (SEQ ID NO:520), Figure 522 (SEQ ID NO:522), Figure 524 (SEQ ID NO:524), Figure 526 (SEQ ID NO:526), Figure 528 (SEQ ID NO:528), Figure 530 (SEQ ID NO:530), Figure 532 (SEQ ID NO:532), Figure 534 (SEQ ID NO:534), Figure 536 (SEQ ID NO:536), Figure 538 (SEQ ID NO:538), Figure 540 (SEQ ID NO:540), Figure 542 (SEQ ID NO:542), Figure 544 (SEQ ID NO:544), Figure 546 (SEQ ID NO:546), Figure 548 (SEQ ID NO:548), Figure 550 (SEQ ID NO:550), Figure 552 (SEQ ID NO:552), Figure 554 (SEQ ID NO:554), Figure 556 (SEQ ID NO:556), Figure 558 (SEQ ID NO:558), Figure 560 (SEQ ID NO:560), Figure 562 (SEQ ID NO:562), Figure 564 (SEQ ID NO:564), Figure 566 (SEQ ID NO:566), Figure 568 (SEQ ID NO:568), Figure 570 (SEQ ID NO:570), Figure 572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), or Figure 610 (SEQ ID NO:610), with its associated signal peptide; or

(c) an amino acid sequence of an extracellular domain of the polypeptide shown in Figure 2 (SEQ ID NO:2), Figure 4 (SEQ ID NO:4), Figure 6 (SEQ ID NO:6), Figure 8 (SEQ ID NO:8), Figure 10 (SEQ ID NO:10), Figure 12 (SEQ ID NO:12), Figure 14 (SEQ ID NO:14), Figure 16 (SEQ ID NO:16), Figure 18 (SEQ ID NO:18), Figure 20 (SEQ ID NO:20), Figure 22 (SEQ ID NO:22), Figure 24 (SEQ ID NO:24), Figure 26 (SEQ ID NO:26), Figure 28 (SEQ ID NO:28), Figure 30 (SEQ ID NO:30), Figure 32 (SEQ ID NO:32), Figure



34 (SEQ ID NO:34), Figure 36 (SEQ ID NO:36), Figure 38 (SEQ ID NO:38), Figure 40 (SEQ ID NO:40),  
Figure 42 (SEQ ID NO:42), Figure 44 (SEQ ID NO:44), Figure 46 (SEQ ID NO:46), Figure 48 (SEQ ID  
NO:48), Figure 50 (SEQ ID NO:50), Figure 52 (SEQ ID NO:52), Figure 54 (SEQ ID NO:54), Figure 56 (SEQ  
ID NO:56), Figure 58 (SEQ ID NO:58), Figure 60 (SEQ ID NO:60), Figure 62 (SEQ ID NO:62), Figure 64  
(SEQ ID NO:64), Figure 66 (SEQ ID NO:66), Figure 68 (SEQ ID NO:68), Figure 70 (SEQ ID NO:70), Figure  
5 72 (SEQ ID NO:72), Figure 74 (SEQ ID NO:74), Figure 76 (SEQ ID NO:76), Figure 78 (SEQ ID NO:78),  
Figure 80 (SEQ ID NO:80), Figure 82 (SEQ ID NO:82), Figure 84 (SEQ ID NO:84), Figure 86 (SEQ ID  
NO:86), Figure 88 (SEQ ID NO:88), Figure 90 (SEQ ID NO:90), Figure 92 (SEQ ID NO:92), Figure 94 (SEQ  
ID NO:94), Figure 96 (SEQ ID NO:96), Figure 98 (SEQ ID NO:98), Figure 100 (SEQ ID NO:100), Figure 102  
(SEQ ID NO:102), Figure 104 (SEQ ID NO:104), Figure 106 (SEQ ID NO:106), Figure 108 (SEQ ID NO:108),  
10 Figure 110 (SEQ ID NO:110), Figure 112 (SEQ ID NO:112), Figure 114 (SEQ ID NO:114), Figure 116 (SEQ  
ID NO:116), Figure 118 (SEQ ID NO:118), Figure 120 (SEQ ID NO:120), Figure 122 (SEQ ID NO:122), Figure  
124 (SEQ ID NO:124), Figure 126 (SEQ ID NO:126), Figure 128 (SEQ ID NO:128), Figure 130 (SEQ ID  
NO:130), Figure 132 (SEQ ID NO:132), Figure 134 (SEQ ID NO:134), Figure 136 (SEQ ID NO:136), Figure  
138 (SEQ ID NO:138), Figure 140 (SEQ ID NO:140), Figure 142 (SEQ ID NO:142), Figure 144 (SEQ ID  
15 NO:144), Figure 146 (SEQ ID NO:146), Figure 148 (SEQ ID NO:148), Figure 150 (SEQ ID NO:150), Figure  
152 (SEQ ID NO:152), Figure 154 (SEQ ID NO:154), Figure 156 (SEQ ID NO:156), Figure 158 (SEQ ID  
NO:158), Figure 160 (SEQ ID NO:160), Figure 162 (SEQ ID NO:162), Figure 164 (SEQ ID NO:164), Figure  
166 (SEQ ID NO:166), Figure 168 (SEQ ID NO:168), Figure 170 (SEQ ID NO:170), Figure 172 (SEQ ID  
NO:172), Figure 174 (SEQ ID NO:174), Figure 176 (SEQ ID NO:176), Figure 178 (SEQ ID NO:178), Figure  
20 180 (SEQ ID NO:180), Figure 182 (SEQ ID NO:182), Figure 184 (SEQ ID NO:184), Figure 186 (SEQ ID  
NO:186), Figure 188 (SEQ ID NO:188), Figure 190 (SEQ ID NO:190), Figure 192 (SEQ ID NO:192), Figure  
194 (SEQ ID NO:194), Figure 196 (SEQ ID NO:196), Figure 198 (SEQ ID NO:198), Figure 200 (SEQ ID  
NO:200), Figure 202 (SEQ ID NO:202), Figure 204 (SEQ ID NO:204), Figure 206 (SEQ ID NO:206), Figure  
208 (SEQ ID NO:208), Figure 210 (SEQ ID NO:210), Figure 212 (SEQ ID NO:212), Figure 214 (SEQ ID  
25 NO:214), Figure 216 (SEQ ID NO:216), Figure 218 (SEQ ID NO:218), Figure 220 (SEQ ID NO:220), Figure  
222 (SEQ ID NO:222), Figure 224 (SEQ ID NO:224), Figure 226 (SEQ ID NO:226), Figure 228 (SEQ ID  
NO:228), Figure 230 (SEQ ID NO:230), Figure 232 (SEQ ID NO:232), Figure 234 (SEQ ID NO:234), Figure  
236 (SEQ ID NO:236), Figure 238 (SEQ ID NO:238), Figure 240 (SEQ ID NO:240), Figure 242 (SEQ ID  
NO:242), Figure 244 (SEQ ID NO:244), Figure 246 (SEQ ID NO:246), Figure 248 (SEQ ID NO:248), Figure  
30 250 (SEQ ID NO:250), Figure 252 (SEQ ID NO:252), Figure 254 (SEQ ID NO:254), Figure 256 (SEQ ID  
NO:256), Figure 258 (SEQ ID NO:258), Figure 260 (SEQ ID NO:260), Figure 262 (SEQ ID NO:262), Figure  
264 (SEQ ID NO:264), Figure 266 (SEQ ID NO:266), Figure 268 (SEQ ID NO:268), Figure 270 (SEQ ID  
NO:270), Figure 272 (SEQ ID NO:272), Figure 274 (SEQ ID NO:274), Figure 276 (SEQ ID NO:276), Figure  
278 (SEQ ID NO:278), Figure 280 (SEQ ID NO:280), Figure 282 (SEQ ID NO:282), Figure 284 (SEQ ID  
35 NO:284), Figure 286 (SEQ ID NO:286), Figure 288 (SEQ ID NO:288), Figure 290 (SEQ ID NO:290), Figure  
292 (SEQ ID NO:292), Figure 294 (SEQ ID NO:294), Figure 296 (SEQ ID NO:296), Figure 298 (SEQ ID  
NO:298), Figure 300 (SEQ ID NO:300), Figure 302 (SEQ ID NO:302), Figure 304 (SEQ ID NO:304), Figure

[illegible]

572 (SEQ ID NO:572), Figure 574 (SEQ ID NO:574), Figure 576 (SEQ ID NO:576), Figure 578 (SEQ ID NO:578), Figure 580 (SEQ ID NO:580), Figure 582 (SEQ ID NO:582), Figure 584 (SEQ ID NO:584), Figure 586 (SEQ ID NO:586), Figure 588 (SEQ ID NO:588), Figure 590 (SEQ ID NO:590), Figure 592 (SEQ ID NO:592), Figure 594 (SEQ ID NO:594), Figure 596 (SEQ ID NO:596), Figure 598 (SEQ ID NO:598), Figure 600 (SEQ ID NO:600), Figure 602 (SEQ ID NO:602), Figure 604 (SEQ ID NO:604), Figure 606 (SEQ ID NO:606), Figure 608 (SEQ ID NO:608), or Figure 610 (SEQ ID NO:610), lacking its associated signal peptide.

20. A method for stimulating the release of TNF- $\alpha$  from human blood, said method comprising contacting said blood with a PRO1079, PRO827, PRO791, PRO1131, PRO1316, PRO1183, PRO1343, PRO1760, PRO1567 or PRO4333 polypeptide, wherein the release of TNF- $\alpha$  from said blood is stimulated.

10

21. A method for stimulating the proliferation or differentiation of chondrocyte cells, said method comprising contacting said cells with a PRO6029 polypeptide, wherein the proliferation or differentiation of said cells is stimulated.

15

22. A method for detecting the presence of tumor in an mammal, said method comprising comparing the level of expression of any PRO polypeptide shown in Table 8 in (a) a test sample of cells taken from said mammal and (b) a control sample of normal cells of the same cell type, wherein a higher level of expression of said PRO polypeptide in the test sample as compared to the control sample is indicative of the presence of tumor in said mammal.

20

23. The method of Claim 22, wherein said tumor is adrenal tumor, lung tumor, colon tumor, breast tumor, prostate tumor, rectal tumor, cervical tumor or liver tumor.

24. An oligonucleotide probe derived from any of the nucleotide sequences shown in the accompanying figures.

25

1/615

**FIGURE 1**

GAAGGCTGCCTCGCTGGTCCGAATTCGGTGGCGCCACGTCCGCCCGTCTCCGCCTTCTGCATCGCGGCTTCGGCG  
GCTTCCACCTAGACACCTAACAGTCGCGGAGCCGGCCGCGTCTGTGAGGGGGTTCGGCACGGGGAGTCGGGCGGTCT  
TGTGCATCTTGGCTACCTGTGGGTCTGAAGATGTCGGACATCGGAGACTGGTTCAGGAGCATCCCCGGCGATCACGC  
GCTATTGGTTCGCCGCCACCGTCCGCCGTGCCCTTGGTCGGCAAACCTCGGCCCTCATCAGCCCGGCTACCTCTTCC  
TCTGGCCCGAAGCCTTCCCTTATCGCTTTCAGATTTGGAGGCCAATCACTGCCACCTTTATTTCCCTGTGGGTC  
CAGGAACCTGGATTTCTTTATTTGGTCAATTTATATTTCTTATATCAGTATTCTACGCGACTTGAAACAGGAGCTT  
TTGATGGGAGGCCAGCAGACTATTTATTCATGCTCCTCTTTAACTGGATTTGCATCGTGATTACTGGCTTAGCAA  
TGGATATGCAGTTGCTGATGATTCCCTCTGATCATGTCAGTACTTTATGTCTGGGCCAGCTGAACAGAGACATGA  
TTGTATCATTTTGGTTTGGAAACACGATTTAAGGCCTGCTATTTACCCTGGGTTATCCTTGGATTCAACTATATCA  
TCGGAGGCTCGGTAATCAATGAGCTTATTGGAAATCTGGTTGGACATCTTTATTTTTTCTAATGTTTCAGATACC  
CAATGGACTTGGGAGGAAGAAATTTCTATCCACACCTCAGTTTTTGTACCGCTGGCTGCCAGTAGGAGAGGAG  
GAGTATCAGGATTTGGTGTGCCCCCTGCTAGCATGAGGCGAGCTGCTGATCAGAATGGCGGAGGCGGAGACACA  
ACTGGGGCCAGGCTTTTCGACTTGGAGACCAGTGAAGGGGCGGCCCTCGGGCAGCCGCTCCTCTCAAGCCACATTT  
CCTCCAGTGTGGGTGCACTTAACAACCTGCGTTCTGGCTAACACTGTTGGACCTGACCCACACTGAATGTAGTC  
TTTCAGTACGAGACAAAGTTTCTTAAATCCCGAAGAAAAATATAAGTGTTCACAAGTTTCACGATTCTCATTCA  
AGTCCTTACTGCTGTGAAGAACAAATACCAACTGTGCAATTTGCAAACTGACTACATTTTTTGGTGTCTTCTCT  
TCTCCCTTTCCGTCTGAATAATGGGTTTTAGCGGGTCTAATCTGCTGGCATTGAGCTGGGCTGGGTACCAA  
ACCTTCCCAAAGGACCTTATCTCTTTCTTGACACATGCCTCTCTCCACTTTTCCCAACCCCCACATTTGCA  
ACTAGAAAAAGTTGCCATAAAATTGCTCTGCCCTTGACAGGTTCTGTTATTTATTGACTTTTGCCAAGGCTGGT  
CACAACAATCATATTCACGTTATTTTCCCTTTTGGTGGCAGAACTGTTACCAATAGGGGGAGAAGACAGCCACG  
GATGAAGCGTTTCTCAGCTTTTGAATTGCTTCGACTGACATCCGTTGTTAACCCTTTGCCACTCTTCAGATATT  
TTTTATAAAAAAGTACCACTGAGTTCATGAGGGCCACAGATTGGTTATTAATGAGATACGAGGGTTGGTGTCTGG  
GTGTTTGTTCCTGAGCTAAGTGATCAAGACTGTAGTGGAGTTGCAGCTAACATGGGTTAGGTTTAAACCATGGG  
GGATGCACCCCTTTGCGTTTCATATGTAGCCCTACTGGCTTTGTGTAGCTGGAGTAGTTGGGTTGCTTTGTGTTA  
GGAGGATCCAGATCATGTTGGCTACAGGGAGATGCTCTCTTTGAGAGGTCCTGGGCATTGATTCCCATTTCAATC  
TCATTCTGGATATGTGTTTCATTGAGTAAAGGAGGAGAGACCCCTCATACGCTATTTAAATGTCACCTTTTTTGCCTA  
TCCCCCGTTTTTGGTTCATGTTTCAATTAATTGTGAGGAAGGCGCAGCTCCTCTCTGCACGTAGATCATTTTTTA  
AAGCTAATGTAAGCACATCTAAGGGAATAACATGATTTAAGGTTGAAATGGCTTTAGAATCATTTGGGTTTGAGG  
GTGTGTTATTTTGAGTCATGAATGTACAAGCTCTGTGAATCAGACCAGCTTAAATACCCACACCTTTTTTTCGTA  
GGTGGGCTTTTCTATCAGAGCTTGGCTCATAACCAAATAAAGTTTTTTGAAGGCCATGGCTTTTCACACAGTTA  
TTTTATTTTATGACGTTATCTGAAAGCAGACTGTTAGGAGCAGTATTGAGTGGCTGTCACACTTTGAGGCAACTA  
AAAAGGCTTCAAACGTTTTGATCAGTTTCTTTTCAGGAAACATTGTGCTCTAACAGTATGACTATTTCTTTCCCC  
ACTCTTAAACAGTGTGATGTGTGTTATCCTAGGAAATGAGAGTTGGCAAACAACCTTCTCATTTTGAATAGAGTTT  
GTGTGTACTTCTCCATATTTAATTTATATGATAAAATAGGTGGGGAGAGTCTGAACCTTAACTGTCATGTTTTGT  
TGTTTCATCTGTGGCCACAATAAAGTTTACTTGTAAATTTTAGAGGCCATTACTCCAATTATGTTGCACGTACAC  
TCATTGTACAGGCGTGGAGACTCATTGTATGTATAAGAATATTTCTGACAGTGAGTGACCCGGAGTCTCTGGTGT  
ACCCTCTTACCAGTCAGCTGCCTGCGAGCAGTCATTTTTCTTAAAGGTTTACAAGTATTTAGAACTTTTAGATT  
CAGGGCAAATGTTTCATGAAGTTATTCTCTTAAACATGGTTAGGAAGCTGATGACGTTATTGATTTTGTCTGGA  
TTATGTTTCTGGAATAATTTTACCAAACAAGCTATTGAGTTTTGACTTGACAAGGCAAACATGACAGTGGAT  
TCTCTTTACAAATGGAAAAAAAATCCTTATTTGTATAAAGGACTTCCCTTTTTTGTAACTAATCCTTTTTTAT  
TGGTAAAAATTGTAAATTAATGTGCAACTTG

2/615

**FIGURE 2**

MSDIGDWFRSIPATRYWFAATVAVPLVGKLGKISPAYLFLWPEAFLYRFQIWRPITATFYFPVGPGTGFLYLVN  
LYFLYQYSTRLETGAFDGRPADYLFMLLFNWICIVITGLAMDMQLLMIPLIMSVLYVWAQLNRDMIVSFWFGTRF  
KACYLPWVILGFNYIIGGSVINELIGNLVGHLYFFLMFRYPMDLGGRNFLSTPQFLYRWLPSRRGGVSGFGVPPA  
SMRRAADQNGGGGRHNWGQGFR LGDQ

**Transmembrane domain:**

amino acids 98-116, 152-172

**N-myristoylation site.**

amino acids 89-95, 168-174, 176-182, 215-221, 221-227, 237-243

**Glycosaminoglycan attachment site.**

amino acids 218-222

3/615

**FIGURE 3**

GAGCGAGGCCGGGGACTGAAGGTGTGGGTGTGAGCCCTCTGGCAGAGGGTTAACCTGGGTCAAATGCACGGATT  
CTCACCTCGTACAGTTACGCTCTCCCGCGGCACGTCCGCGAGGACTTGAAGTCCTGAGCGCTCAAGTTTGTCCGT  
AGGTGAGAGAAGGCCATGGAGGTGCCGCCACCGGCACCGCGAGCTTTCTCTGTAGAGCATTGTGCCTATTTCC  
CCGAGTCTTTGCTGCCGAAGCTGTGACTGCCGATTCCGGAAGTCCTTGAGGAGCGTCAGAAGCGGCTTCCCTACGT  
CCCAGAGCCCTATTACCCGGAATCTGGATGGGACCGCCTCCGGGAGCTGTTTGGCAAAGATGAACAGCAGAGAAT  
TTCAAAGGACCTTGCTAATATCTGTAAGACGGCAGCTACAGCAGGCATCATTGGCTGGGTGTATGGGGGAATACC  
AGCTTTTATTCATGCTAAACAACAATACATTGAGCAGAGCCAGGCAGAAATTTATCATAACCGGTTTGATGCTGT  
GCAATCTGCACATCGTGCTGCCACACGAGGCTTCATTGTTATGGCTGGCGCTGGGGTTGGAGAACTGCAGTGTT  
TGTGACTATATTCAACACAGTGAACACTAGTCTGAATGTATACCGAAATAAAGATGCCTTAAGCCATTTTGTAAAT  
TGCAGGAGCTGTCACGGGAAGTCTTTTAGGATAAACGTAGGCCTGCGTGGCCTGGTGGCTGGTGGCATAATTGG  
AGCCTTGCTGGGCACTCCTGTAGGAGGCCTGCTGATGGCATTTCAGAAGTACGCTGGTGAGACTGTTCAAGGAAAG  
AAAACAGAAGGATCGAAAGGCACTCCATGAGCTAAACTGGAAGAGTGGAAGGCAGACTACAAGTTACTGAGCA  
CCTCCCTGAGAAAATTGAAAGTAGTTTACGGGAAGATGAACCTGAGAATGATGCTAAGAAAATTGAAGCACTGCT  
AAACCTTCCTAGAAACCCTTCAGTAATAGATAAACAAGACAAGGACTGAAGAGTCTCTGAACCTGAAACTCACTG  
GAGAGCTGAAGGGAGCTGCCATGTCCGATGAATGCCAACAGACAGGCCACTCTTTGGTCAGCCTGCTGACAAATT  
TAAGTGCTGGTACCTGTGGTGGCAGTGGCTTGCTCTTGTCTTTTCTTTTCTTTTAACTAAGAATGGGGCTGTT  
GTACTCTCACTTTACTTATCCTTAAATTTAAATACATACTTATGTTTGTATTAATCTATCAATATATGCATACAT  
GGATATATCCACCCACCTAGATTTTAAGCAGTAAATAAACATTTGCAAAAGATTAAAGTTGAATTTTACAGTTT

4/615

**FIGURE 4**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA23318
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32190, pI: 9.03, NX(S/T): 2
MEVPPPAPRSFLCRALCLFPRVFAAEAVTADSEVLEERQKRLPYVPEPYYPESGWDRLRELF GKDEQQRISKDLA
NICKTAATAGIIGWVYGGIPAFIHAKQQYIEQSQAETIYHNRFDAVQSAHRAATRGFIRYGWRWGWRTAVFVTIFN
TVNTSLNVYRNKDALSHEFVIAGAVTGSLFRINVGLRGLVAGGIIGALLGTPVGGLLMAFQKYAGETVQERKQKDR
KALHELKLEEWKGRLQVTEHLPEKIESSLREDEPENDAKKIEALLNLPRNPSVIDKQDKD
```

**Important Features:****Signal Peptide:**

amino acids 1-24

**Transmembrane domains:**

amino acids 76-96 and 171-195

**N-glycosylation site:**

amino acids 153-156

5/615

**FIGURE 5**

CGGACGCGTGGGCGCGGGACGCCGGCAGGGTTGTGGCGCAGCAGTCTCCTTCCTGCGCGCGCGCCTGAAGTCGGC  
GTGGGCGTTTTGAGGAAGCTGGGATACAGCATTTAATGAAAAATTTATGCTTAAGAAGTAAAAATGGCAGGCTTCC  
TAGATAATTTTCGTTGGCCAGAATGTGAATGTATTGACTGGAGTGAGAGAAGAAATGCTGTGGCATCTGTTGTCTG  
CAGGTATATTGTTTTTACAGGCTGGTGGATAATGATTGATGCAGCTGTGGTGTATCCTAAGCCAGAACAGTTGA  
ACCATGCCTTTCACACATGTGGTGTATTTCCACATTGGCTTTCTTCATGATAAATGCTGTATCCAATGCTCAGG  
TGAGAGGTGATAGCTATGAAAGCGGCTGTTTAGGAAGAACAGGTGCTCGAGTTTGGCTTTTCATTGGTTTCATGT  
TGATGTTTGGGTCACTTATTGCTTCCATGTGGATTCTTTTTGGTGCATATGTTACCCAAAATACTGATGTTTATC  
CGGGACTAGCTGTGTTTTTCAAATGCACTTATATTTTTTAGCACTCTGATCTACAAATTTGGAAGAACCGAAG  
AGCTATGGACCTGAGATCACTTCTTAAGTCACATTTTCCTTTTGTATATTCTGTTTGTAGATAGGTTTTTTATC  
TCTCAGTACACATTGCCAAATGGAGTAGATTGTACATTAAATGTTTTGTTTCTTTACATTTTTATGTTCTGAGTT  
TTGAAATAGTTTTATGAAATTTCTTTATTTTTCATTCATAGACTGTTAATATGTATATAATACAAGACTATATG  
AATTGGATAATGAGTATCAGTTTTTTATTCCTGAGATTTAGAACTTGATCTACTCCCTGAGCCAGGGTTACATCA  
TCTTGTCATTTTAGAAAGTAACCACTCTTGTCTCTCTGGCTGGGCACGGTGGCTCATGCCTGTAATCCCAGCACTT  
TGGGAGGCCGAGGCCGGCCGATTGCTTGAGGTCAAGTGTGAGACCAGCCTGGCCAACATGGCGAAAACCCCATC  
TACTAAAAATACAAAAATTAGCCAGGCATGGTGGTGGGTGCCTGTAATCCCAGCTACCTGGGAGGCTGAGGCAGG  
AGAATCGCTTGAACCCGGGGGGCAGAGGTTGCAGTGAGCTGAGTTTGCGCCACTGCACTCTAGCCTGGGGGAGAA  
AGTGAAACTCCCTCTCAAAAAAAGACCACTCTCAGTATCTCTGATTTCTGAAGATGTACAAAAAATATAGCTT  
CATATATCTGGAATGAGCACTGAGCCATAAAAGGTTTTAGCAAGTTGTAACCTATTTTGGCCTAAAAATGAGGT  
TTTTTTGGTAAAGAAAAAATATTTGTTCTTATGTATTGAAGAAGTGACTTTTATATAATGATTTTTTAAATGCC  
CAAAGGACTAGTTTGAAAGCTTCTTTTAAAAAGAATTCCTCTAATATGACTTTATGTGAGAA



6/615

## **FIGURE 6**

MAGFLDNFRWPECECIDWSERRNAVASVVAGILFFTGWIMIDAADVYPKPEQLNHAFHTCGVFSTLAFMINAV  
SNAQVRGDSYESGCLGRTGARVWLFIFGMLMFGSLIASMWILFGAYVTQNTDVYPGLAVFFQNALIFFSTLIYKF  
GRTEELWT

**Important features:**

**Signal peptide:**

amino acids 1-44

**Transmembrane domains:**

amino acids 23-42 (type II), 60-80, 97-117, 128-148

7/615

**FIGURE 7**

CGGTGGTTTTTGTCTGCAATAGGCGGCTTAGAGGGAGGGGCTTTTTCGCCTATACCTACTGTAGCTTCTCCAGG  
TATGGACCCTAAAGGCTACTGCTGCTACTACGGGGCTAGACAGTTACTGTCTCAGCTCTAGGATGTGCGTCTCTTC  
CACTAGAAGCTCTTCTGAGGGAGGTAATTAACCAAGTGGAAATGGAAAAACAGTGCTGTAGTCATCCTGTAATA  
TGCTCCTTGTCAACAATGTATACATTCTGCTAGGTGCCATATTCATTGCTTTAAGCTCAAGTCGCATCTTACTA  
GTGAAGTATTCTGCCAATGAAGAAAAACAAGTATGATTATCTTCCAACCTACTGTGAATGTGTGCTCAGAAGCTGGTG  
AAGCTAGTTTTCTGTGTGCTTGTGTCTTCTGTCTGTTATATAAGAAAGATCATCAAAGTAGAAATTTGAAATATGCT  
TCCTGGAAGGAATCTCTGATTTTCATGAAGTGGTCCATTCTCGCCTTTCTTTATTTCTGGATAACTTGATTGTC  
TTCTATGTCTGTCTCTATCTTCAACCAGCCATGGCTGTTATCTTCTCAAATTTTAGCATTATAACAACAGCTCTT  
CTATTTCAGGATAGTGCTGAAGAGGCGTCTAAACTGGATCCAGTGGGCTTCCCTCCTGACTTTATTTTTGTCTATT  
GTGGCCTTGACTGCCGGGACTAAAACTTTACAGCACAACTTGGCAGGACGTGGATTTTCATCAGGATGCCCTTTTTC  
AGCCCTTCCAATTCCTGCCTTCTTTTCAGAAGTGAGTGTCCAGAAAAAGACAATTGTACAGCAAAGGAATGGACT  
TTTCCTGAAGCTAAATGGAACACCACAGCCAGAGTTTTCAGTCACATCCGTCTTGGCATGGGCCATGTTCTTATT  
ATAGTCCAGTGTTTTATTTCTTCAATGGCTAATATCTATAATGAAAAGATACTGAAGGAGGGGAACAGCTCACT  
GAAAGCATCTTCATACAGAACAGCAAACTCTATTTCTTTGGCATTCTGTTTAATGGGCTGACTCTGGGCCTTCAG  
AGGAGTAACCGTGATCAGATTAAAGAACTGTGGATTTTTTATGGCCACAGTGCATTTTCAGTAGCCCTTATTTTT  
GTAAGTGCATTCAGGGCCTTTTCAGTGGCTTTTCATTCTGAAGTTCCTGGATAACATGTTCCATGTCTTGATGGCC  
CAGGTTACCACTGTCTATTATCACAACTGTCTGTCTGGTCTTTGACTTCAGGCCCTCCCTGGAATTTTCTTTG  
GAAGCCCCATCAGTCCCTTCTCTATATTTATTTATAATGCCAGCAAGCCTCAAGTTCCGGAATACGCCACCTTAGG  
CAAGAAAGGATCCGAGATCTAAGTGGCAATCTTTGGGAGCGTTCCAGTGGGGATGGAGAAGAACTAGAAAAGACTT  
ACCAAACCAAGAGTGATGAGTCAGATGAAGATACTTTCTAACTGGTACCCACATAGTTTGCAGCTCTCTTGAAC  
CTTATTTTTCATTTTTCAGTGTGTAATATTTATCTTTTCACTTTGATAAACCAGAAATGTTTCTAAATCCTAA  
TATTCTTTGCATATATCTAGCTACTCCCTAAATGGTTCATCCAGGCTTAGAGTACCAAGGCTAAGAAATTC  
TAAAGAACTGATACAGGAGTAACAATATGAAGAATTCATTAATATCTCAGTACTTGATAAATCAGAAAGTTATAT  
GTGCAGATTATTTTCTTGGCCTTCAAGCTTCCAAAAAACTTGTAATAATCATGTTAGCTATAGCTTGTATATAC  
ACATAGAGATCAATTTGCCAAATATTCACAATCATGTAGTTCTAGTTTACATGCCAAAGTCTTCCCTTTTAAACA  
TTATAAAAGCTAGGTTGTCTCTTGAATTTTGGGCCCCTAGAGATAGTCATTTTGCAAGTAAAGAGCAACGGGACC  
CTTTCTAAAAACGTTGGTTGAAGGACCTAAATACCTGGCCATACCATAGATTTGGGATGATGTAGTCTGTGCTAA  
ATATTTTGTCTGAAGAAGCAGTTTCTCAGACACAACATCTCAGAATTTTAAATTTTAAAGAAATTCATGGGAAATTTGG  
ATTTTTGTAATAATCTTTTGATGTTTTAAACATTGGTTCCCTAGTCACCATAGTTTACCATTGTATTTTAAAGTCA  
TTTAAACAAGCCACGGTGGGGCTTTTTCTCCTCAGTTTGAGGAGAAAAATCTTGATGTCTTACTCCTGAATTA  
TTACATTTTGGAGAATAAGAGGGCATTTTATTTTATTAGTTACTAATTCAGCTGTGACTATTGTATATCTTTCC  
AAGAGTTGAAATGCTGGCTTCAGAATCATACCAGATTGTCTAGTGAAGCTGATGCCTAGGAACCTTTTAAAGGGATC  
CTTTCAAAGGATCACTTAGCAAACACATGTTGACTTTTAACTGATGTATGAATATTAATACTCTAAAAATAGAA  
AGACCAGTAATATATAAGTCACTTTACAGTGCTACTTCACACTTAAAAGTGCATGGTATTTTTCATGGTATTTTG  
CATGCAGCCAGTTAACTCTCGTAGATAGAGAAGTCAGGTGATAGATGATATTAAAAATTAGCAAACAAAAGTGAC  
TTGCTCAGGGTCATGCAGCTGGGTGATGATAGAAGAGTGGGCTTTAACTGGCAGGCCTGTATGTTTACAGACTAC  
CATACTGTAATATAGGCTTTATGGTGTCTATTCTCAGAACTTATACATTTCTGCTCTCCTTTCTCCTAAGTTTC  
ATGCAGATGAATATAAGGTAATATACTATTATATAATTCATTTGTGATATCCACAATAATATGACTGGCAAGAAT  
TGGTGGAAATTTGTAATTAATAATTTATTAACCT

8/615

**FIGURE 8**

MEKQCCSHPVICSLSTMYTFLLGAIFIALSSSRILLVKYSANEENKYDYLPTTVNVCSELVKLVFCVLVSFCVIK  
KDHQSRNLKYASWKEFSDFMKWSIPAFLYFLDNLIVFYVLSYLQPAMAVIFS NFSIITALLFRIVLKRRRLNWIQ  
WASLLTLFLSIVALTAGTKTLQHNLAGRGFHHDAFFSPSNSCLLFRSECPRKDNCTAKEWTFPEAKWNTTARVFS  
HIRLGMGHVLIIVQCFISSMANIYNEKILKEGNQLTESIFIQNSKLYFFGILFNGLTLGLQRSNRDQIKNCGFFY  
GHSAFSVALIFVTAFQGLSVAFILKFLDNMFHVLMAQVTTVIITTVSVLVFDFRPSLEFFLEAPSVLLSIFIYNA  
SKPQVP EYAPRQERIRDL SGNLWERS SSGDGEELERLTKPKSDESD EDTF

**Transmembrane domains:**

amino acids 16-36 (type II), 50-74, 147-168, 229-250, 271-293, 298-318,  
328-368

**N-glycosylation sites.**

amino acids 128-132, 204-208, 218-222, 374-378

**Glycosaminoglycan attachment site.**

amino acids 402-406

**N-myristoylation sites.**

amino acids 257-263, 275-281, 280-286, 284-290, 317-323

9/615

**FIGURE 9**

GGGGCTTCGGCGCCAGCGGCCAGCGCTAGTCGGTCTGGTAAGGATTTACAAAAGGTGCAGGTATGAGCAGGTCTG  
AAGACTAACATTTTGTGAAGTTGTAAAACAGAAAACCTGTTAGAAATGTGGTGGTTTCAGCAAGGCCTCAGTTTC  
CTTCCTTCAGCCCTTGTAATTTGGACATCTGCTGCTTCATATTTTCATACATTACTGCAGTAACACTCCACCAT  
ATAGACCCGGCTTTACCTTATATCAGTGACACTGGTACAGTAGCTCCAGAAAAATGCTTATTTGGGGCAATGCTA  
AATATTGCGGCAGTTTATGCATTGCTACCATTTATGTTTCGTTATAAGCAAGTTCATGCTCTGAGTCCTGAAGAG  
AACGTTATCATCAAATTAACAAGGCTGGCCTTGACTTGGAATACTGAGTTGTTTAGGACTTTCTATTGTGGCA  
AACTTCCAGAAAACAACCCTTTTGTCTGCACATGTAAGTGGAGCTGTGCTTACCTTTGGTATGGGCTCATTATAT  
ATGTTTGTTCAGACCATCCTTTCCTACCAAATGCAGCCCAAATCCATGGCAAACAAGTCTTCTGGATCAGACTG  
TTGTTGGTTATCTGGTGTGGAGTAAGTGCACCTTAGCATGCTGACTTGCTCATCAGTTTGCACAGTGGCAATTT  
GGGACTGATTTAGAACAGAACTCCATTGGAACCCCGAGGACAAAGGTTATGTGCTTCACATGATCACTACTGCA  
GCAGAATGGTCTATGTCATTTTCCTTCTTTGGTTTTTTCCTGACTTACATTCGTGATTTTCAGAAAAATTTCTTTA  
CGGGTGGAAGCCAATTTACATGGATTAACCCTCTATGACACTGCACCTTGCCCTATTAACAATGAACGAACACGG  
CTACTTTCAGAGATATTTGATGAAAGGATAAAATATTTCTGTAATGATTATGATTCTCAGGGATTGGGGAAAGG  
TTCACAGAAGTTGCTTATTTCTTCTCTGAAATTTTCAACCACTTAATCAAGGCTGACAGTAACACTGATGAATGCT  
GATAATCAGGAAACATGAAAGAAGCCATTTGATAGATTATTCTAAAGGATATCATCAAGAAGACTATTA AAAACA  
CCTATGCCTATACTTTTTTATCTCAGAAAATAAAGTCAAAAGACTATG

10/615

**FIGURE 10**

MWWFQQGLSFLPSALVIWTSAAFIIFSITAVTLHHIDPALPYISDTGTVAPEKCLFGAMLNIAAVLCIATIIYVRY  
KQVHALSPEENVIIKLNKAGLVLGILSCLGLSIVANFQKTTLFAAHVSGAVLTFGMGSLYMFVQTILSYQMOPKI  
HGKQVFWIRLLLVWCGVSALSMLTCSSVLHSGNFGTDLEQKLHWNPEDKGYVLHMITTAAEWSMSFSFFGFFLT  
YIRDFQKISLRVEANLHGLTLYDTAPCPINNERTRLLSRDI

11/615

**FIGURE 11**

CCCACGCGTCCGCCCCCGCTGCGTCCCGGAGTGCAAGTGAGCTTCTCGGCTGCCCCGCGGGCCGGGGTGCGGAG  
CCGACATGCGCCCCGCTTCTCGGCCTCCTTCTGGTCTTCGCCGGGTGCACCTTCGCCTTGACTTGCTGTCGACGC  
GACTGCCCCGCGGGCGGAGACTGGGCTCCACCGAGGAGGCTGGAGGCAGGTGCTGTGGTCCCCCTCCGACCTGG  
CAGAGCTGCGGGAGCTCTCTGAGGTCCCTTCGAGAGTACCGGAAGGAGCACCAGGCCTACGTGTTCCCTGCTCTTCT  
GCGGCGCCTACCTCTACAAACAGGGCTTTGCCATCCCCGGCTCCAGCTTCCTGAATGTTTTAGCTGGTGCCTTGT  
TTGGGCCATGGCTGGGGCTTCTGCTGTGCTGTGTGTGACCTCGGTGGGTGCCACATGCTGCTACCTGCTCTCCA  
GTATTTTTGGCAAACAGTTGGTGGTGTCTACTTTCTTGATAAAAGTGGCCCTGCTGCAGAGAAAGGTGGAGGAGA  
ACAGAAACAGCTTGTTTTTTTTCTTATTGTTTTTGAGACTTTTCCCCATGACACCAAACCTGGTTCTTGAACCTCT  
CGGCCCCAATTCTGAACATTCCCATCGTGCAGTTCTTCTCTCAGTTCTTATCGGTTTGATCCCATATAATTTCA  
TCTGTGTGCAGACAGGGTCCATCCTGTCAACCCTAACCTCTCTGGATGCTCTTTCTCCTGGGACACTGTCTTTA  
AGCTGTTGGCCATTGCCATGGTGGCATTAAATTCCTGGAACCCTCATTAAAAAATTTAGTCAGAAACATCTGCAAT  
TGAATGAAACAAGTACTGCTAATCATATACACAGTAGAAAAGACACATGATGATCTGGATTTTCTGTTTGCCACATCC  
CTGGACTCAGTTGCTTATTTGTGTAATGGATGTGGTCCCTCTAAAGCCCCTCATTGTTTTTGATTGCCTTCTATAG  
GTGATGTGGACACTGTGCATCAATGTGCAGTGTCTTTTCAGAAAGGACACTCTGCTCTTGAAGGTGTATTACATC  
AGGTTTTCAAACCAGCCCTGGTGTAGCAGACACTGCAACAGATGCCTCCTAGAAAATGCTGTTTGTGGCCGGGCG  
CGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCCGAGGCCGGTGATTCAACAAGGTCAGGAGTTCAAGACC  
AGCCTGGCCAAGATGGTGAAATCCTGTCTCTAATAAAAAATACAAAAATTAGCCAGGCGTGGTGGCAGGCACCTGT  
AATCCCAGCTACTCGGGAGGCTGAGGCAGGAGAATTGCTTGAACCAAGGTGGCAGAGGTTGCAGTAAGCCAAGAT  
CACACCACTGCACTCCAGCCTGGGTGATAGAGTGAGACACTGTCTTGAC

12/615

**FIGURE 12**

MRPLLGLLLVFAGCTFALYLLSTRLPGRRLGSTEEAGGRSLWFPSDLAELRESEVLREYRKEHQAYVFLLFCG  
AYLYKQGFAIPGSSFLNVLGALFGPWLGLLLCCVLTSVGATCCYLLSSIFGKQLVVSYPDKVALLQRKVEENR  
NSLFFFLLFLRLFPMTPNWFLNLSAPILNIPVQFFSVLIGLIPYNFICVQTGSILSTLTSLDALFSWDTVEKL  
LAIAMVALIPGTLIKKEFSQKHLQLNETSTANHIHSRKDT

**Important features:****Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 101-123, 189-211

**N-glycosylation sites.**

amino acids 172-176, 250-254

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 240-244, 261-265

**N-myristoylation site.**

amino acids 13-19, 104-110, 115-121, 204-210

**Amidation site.**

amino acids 27-31

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 4-15

**Protein splicing proteins.**

amino acids 25-31

**Sugar transport proteins.**

amino acids 162-172

13/615

**FIGURE 13**

CGGACGCGTGGGCGGACGCGTGGGGGAGAGCCGCAGTCCCGGCTGCAGCACCTGGGAGAAGGCAGACCGTGTGAG  
GGGGCCTGTGGCCCCAGCGTGCTGTGGCCTCGGGGAGTGGGAAGTGGAGGCAGGAGCCTTCCTTACACTTCGCCA  
**TG**AGTTTTCTCATCGACTCCAGCATCATGATTACCTCCCAGATACTATTTTTTGGATTTGGGTGGCTTTTCTTCA  
TGCGCCAATTGTTTAAAGACTATGAGATACGTGAGTATGTTGTACAGGTGATCTTCTCCGTGACGTTTGCATTTT  
CTTGCACCATGTTTGAGCTCATCATCTTTGAAATCTTAGGAGTATTGAATAGCAGCTCCCGTTATTTTCACTGGA  
AAATGAACCTGTGTGTAATTCTGCTGATCCTGGTTTTTCATGGTGCCTTTTACATTGGCTATTTTATTGTGAGCA  
ATATCCGACTACTGCATAAAACAACGACTGCTTTTTTCTGTCTCTTATGGCTGACCTTTATGTATTTCTTCTGGA  
AACTAGGAGATCCCTTTCCCATCTCAGCCCAAAACATGGGATCTTATCCATAGAACAGCTCATCAGCCGGGTTG  
GTGTGATTGGAGTGACTCTCATGGCTCTCTTTCTGGATTTGGTGCTGTCAACTGCCCATACACTTACATGTCTT  
ACTTCCTCAGGAATGTGACTGACACGGATATTCTAGCCCTGGAACGGCGACTGCTGCAAACCATGGATATGATCA  
TAAGCAAAAAGAAAAGGATGGCAATGGCACGGAGAACAATGTTCCAGAAGGGGGAAGTGCATAACAAACCATCAG  
GTTTCTGGGGAATGATAAAAAGTGTTACCACTTCAGCATCAGGAAGTGAAAATCTTACTCTTATTCAACAGGAAG  
TGGATGCTTTGGAAGAATTAAGCAGGCAGCTTTTTCTGGAAACAGCTGATCTATATGCTACCAAGGAGAGAATAG  
AATACTCCAAAACCTTCAAGGGGAAATATTTTAAATTTTCTTGGTTACTTTTTCTCTATTTACTGTGTTTGGAAAA  
TTTTTCATGGCTACCATCAATATTGTTTTTGATCGAGTTGGGAAAACGGATCCTGTCAAGAGGCATTGAGATCA  
CTGTGAATTATCTGGGAATCCAATTTGATGTGAAGTTTTGGTCCCAACACATTTCTTCTATTCTTGTGGGAATAA  
TCATCGTCACATCCATCAGAGGATTGCTGATCACTCTTACCAAGTTCTTTTATGCCATCTCTAGCAGTAAGTCCT  
CCAATGTCATTGTCTCTGCTATTAGCACAGATAATGGGCATGTACTTTGTCTCCTCTGTGCTGCTGATCCGAATGA  
GTATGCCTTTAGAATACCGCACCATTAATCACTGAAGTCCTTGGAGAACTGCAGTTCAACTTCTATCACCGTTGGT  
TTGATGTGATCTTCTGGTCAGCGCTCTCTCTAGCATACTCTTCTCTATTTGGCTCACAAACAGGCACCAGAGA  
AGCAAATGGCACCT**TGA**ACTTAAGCCTACTACAGACTGTTAGAGGCCAGTGGTTTCAAATTTAGATATAAGAGG  
GGGGAAAAATGGAACCAGGGCCTGACATTTTATAAACAAACAAAATGCTATGGTAGCATTTTTCACCTTCATAGC  
ATACTCCTTCCCCGTGAGGTGATACTATGACCATGAGTAGCATCAGCCAGAACATGAGAGGGAGAACCTAACTCAA  
GACAATACTCAGCAGAGAGCATCCCGTGTGGATATGAGGCTGGTGTAGAGGCGGAGAGGAGCCAAGAACTAAAG  
GTGAAAAATACACTGGAACCTCTGGGGCAAGACATGTCTATGGTAGCTGAGCCAAACACGTAGGATTTCCGTTTTA  
AGGTTACATGGAAGGTTATAGCTTTGCCTTGAGATTGACTCATTAATCAGAGACTGTAACAAAAA  
AAAAAAGGGCGCGCGACTCTAGAGTCGACCTGCAGAAGCTTGGCCGCCATGGCCCACTTGTTTATTG  
CAGCTTATAATG



14/615

**FIGURE 14**

MSFLIDSSIMITSQILFFGFGWLFFMRQLFKDYEIRQYVVQVIFSVTFAFSCTMFELIIFEILGVLNSSSRYPFH  
KMNL CVILLILVFMVFPFYIGYFIVSNIRLLHKQRLLFSCLLWLTFMYFFWKLGDFFPILSPKHGILSIEQLISRV  
GVIGVTLMALLSGFGAVNCPYTYMSYFLRNVTDTDILALERRLLQTMDMIISKKKRMAMARRTMFQKGEVHNKPS  
GFWGMIKSVTTSASGSENLTLIQQEVDLEELSRQLFLETADLYATKERIEYSKTFKGKYNFLGYFFSIYCVWK  
IFMATINIVFDRVGKTDVPTRGIEITVNYLGIQFDVKFWSQHISFILVGIIIVTSIRGLLITLTKFFYAISSSKS  
SNVIVLLLAQIMGMYFVSSVLLIRMSMPLEYRTIITEVLGELQFNFYHRWFDVIFLVSALSSILFLYLAKQAPE  
KQMAP

**Important features:****Signal peptide:**

amino acids 1-23

**Potential transmembrane domains:**

amino acids 37-55, 81-102, 150-168, 288-311, 338-356, 375-398, 425-444

**N-glycosylation sites.**

amino acids 67-70, 180-183 and 243-246

**Eukaryotic cobalamin-binding proteins**

amino acids 151-160

15/615

**FIGURE 15**

GACGGAAGAACAGCGCTCCCGAGGCCGCGGGAGCCTGCAGAGAGGACAGCCGGCCTGCGCCGGGACATGCGGCCC  
CAGGAGCTCCCCAGGCTCGCGTTCCCGTTGCTGCTGTTGCTGTTGCTGCTGCTGCCGCCGCCGCGTGCCTGCC  
CACAGCGCCACGCGCTTCGACCCACCTGGGAGTCCCTGGACGCCGCCAGCTGCCCGCGTGGTTTGACCAGGCC  
AAGTTCGGCATCTTCATCCACTGGGGAGTGTTTTCCGTGCCAGCTTCGGTAGCGAGTGTTCTGGTGGTATTGG  
CAAAAGGAAAAGATACCGAAGTATGTGGAATTTATGAAAGATAATTACCCTCCTAGTTTCAAATATGAAGATTTT  
GGACCACTATTTACAGCAAAATTTTTTAATGCCAACCAGTGGGCAGATATTTTTCAGGCCTCTGGTGCCAAATAC  
ATTGTCTTAACCTCCAAACATCATGAAGGCTTTACCTTGTGGGGTCAAGAATATTCGTGGAACTGGAATGCCATA  
GATGAGGGGCCCCAAGAGGGACATTGTCAAGGAACCTGAGGTAGCCATTAGGAACAGAACTGACCTGCGTTTTGGA  
CTGTACTATTCCTTTTTGAATGGTTTCATCCGCTCTTCTTGAGGATGAATCCAGTTCATTCCATAAGCGGCAA  
TTTCCAGTTTCTAAGACATTGCCAGAGCTCTATGAGTTAGTGAACAACTATCAGCCTGAGGTTCTGTGGTCCGAT  
GGTGACGGAGGAGCACCGGATCAATACTGGAACAGCACAGGCTTCTTGGCCTGGTTATATAATGAAAGCCAGTT  
CGGGGCACAGTAGTCACCAATGATCGTTGGGGAGCTGGTAGCATCTGTAAGCATGGTGGCTTCTATACCTGCAGT  
GATCGTTATAACCCAGGACATCTTTGCCACATAAATGGGAAAACCTGCATGACAATAGACAAACTGTCTGGGGC  
TATAGGAGGGAAGCTGGAATCTCTGACTATCTTACAATTGAAGAATTGGTGAAGCAACTGTAGAGACAGTTTCA  
TGTGGAGGAAATCTTTGATGAATATTGGGCCCACACTAGATGGCACCATTCTGTAGTTTTTGAGGAGCGACTG  
AGGCAAGTGGGGTCTGGCTAAAAGTCAATGGAGAAGCTATTTATGAAACCTATACCTGGCGATCCCAGAATGAC  
ACTGTCACCCCAGATGTGTGGTACACATCCAAGCCTAAAGAAAAATTAGTCTATGCCATTTTTCTTAAATGGCCC  
ACATCAGGACAGCTGTTCTTGGCCATCCCAAAGCTATTCTGGGGGCAACAGAGGTGAACTACTGGGCCATGGA  
CAGCCACTTAACTGGATTTCTTGGAGCAAAATGGCATTATGGTAGAACTGCCACAGCTAACCATTCATCAGATG  
CCGTGTAAATGGGGCTGGGCTCTAGCCCTAACTAATGTGATCTAAAGTGCAGCAGAGTGGCTGATGCTGCAAGTT  
ATGTCTAAGGCTAGGAACATACAGGTGTCTATAATTGTAGCACATGGAGAAAGCAATGTAACTGGATAAGAAAA  
TTATTTGGCAGTTCAGCCCTTTCCCTTTTTCCCACTAAATTTTTCTTAAATTACCCATGTAAACCATTTTAACTCT  
CCAGTGCACCTTTGCCATTAAAGTCTCTTCACATTGATTTGTTTCCATGTGTGACTCAGAGGTGAGAATTTTTTCA  
CATTATAGTAGCAAGGAATTGGTGGTATTATGGACCGAACTGAAAATTTTATGTTGAAGCCATATCCCCATGAT  
TATATAGTTATGCATCACTTAATATGGGGATATTTTCTGGGAAATGCATTGCTAGTCAATTTTTTTTTGTGCCAA  
CATCATAGAGTGTATTTACAAAATCCTAGATGGCATAGCCTACTACACACCTAATGTGTATGGTATAGACTGTTG  
CTCCTAGGCTACAGACATATACAGCATGTTACTGAATACTGTAGGCAATAGTAACAGTGGTATTTGTATATCGAA  
ACATATGGAACATAGAGAAGGTACAGTAAAAATACTGTAAATAAATGGTGACCTGTATAGGGCACTTACCAC  
GAATGGAGCTTACAGGACTGGAAGTTGCTCTGGGTGAGTCAGTGAGTGAATGTGAAGGCCTAGGACATTATTGAA  
CACTGCCAGACGTTATAAAATACTGTATGCTTAGGCTACACTACATTTATAAAAAAAGTTTTTCTTCTTCAATT  
ATAAATTAACATAAGTGTACTGTAACCTTTACAAACGTTTTAATTTTTTAAACCTTTTTGGCTCTTTGTAATAAC  
ACTTAGCTTAAACATAAACTCATTGTGCAATGTAA

16/615

**FIGURE 16**

MRPQELPRLAFPLLLLLLLLLPPPPCPAHSATRFDPWTWESLDARQLPAWFDQAKFGIFIHGVSFVPSFGSEWFW  
WYWQKEKIPKYVEFMKDNYPSPFKYEDFGPLFTAKFFNANQWADIFQASGAKYIVLTSKHHEGFTLWGSEYSWNW  
NAIDEGPKRDIVKELEVAIRNRTDLRFGLYYSLFEWFHPLFLEDESSSFHKRQFPVSKTLPELYELVNNYQPEVL  
WSDGDGGAPDQYWNSTGFLAWLYNESPVRGTVVTNDRWGAGSICKHGGFYTCSDRYNPGHLLPHKWENCMTIDKL  
SWGyrREAGISDYLTIIEELVKQLVETVSCGNNLLMNIGPTLDGTISVVFEERLRQVGSWLKVNGEAIYETYTWRS  
QNDTVTPDVWYTSKPKEKLVYAIFLKWPTSGQLFLGHPKAILGATEVKLLGHGQPLNWISLEQNGIMVELPQITI  
HQMPCCKWGWALALTNVI

**Signal sequence:**

amino acids 1-28

**N-glycosylation site.**

amino acids 171-175, 239-243, 377-381

**Casein kinase II phosphorylation site.**

amino acids 32-36, 182-186, 209-213, 227-231, 276-280, 315-319, 375-375

**Tyrosine kinase phosphorylation site.**

amino acids 361-369, 389-397

**N-myristoylation site.**

amino acids 143-149, 178-184, 255-261, 272-278, 428-434

**Leucine zipper pattern.**

amino acids 410-432

**Alpha-L-fucosidase putative active site.**

amino acids 283-295

**FIGURE 17**

[illegible]

18/615

**FIGURE 18**

MKFLLDILLLLPLLIIVCSLESFVKLFIPKRRKSVTGEIVLITGAGHGIGRLTAYEFAKLKSKL  
VLWDINKHGLEETAACKCKGLGAKVHTFVVDSCNREDIYSSAKKVKAIEIGDVSILVNNAGVVYT  
SDLFATQDPQIEKTFEVNVLAHFWTTKAFLPAMTKNNHGHIVTVASAAGHVSVPFLLAYCSSK  
FAAVGFHKTLTDELAALQITGVKTTCLCPNFVNTGFIKNPSTSLGPTLEPEEVVNRMLMHGILT  
EQKMIFIPSSIAFLTTLERILPERFLAVLKRKISVKFDAVIGYKMQ

**Signal sequence:**

amino acids 1-19

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-34, 283-287

**Casein kinase II phosphorylation site.**

amino acids 52-56, 95-99, 198-202, 267-271

**N-myristoylation site.**

amino acids 43-49, 72-78, 122-128, 210-216

19/615

**FIGURE 19**

CCCACGCGTCCGCTCCGCGCCCTCCCCCGCCTCCCGTGCGGTCCGTCGGTGGCCTAGAGAT  
GCTGCTGCCGCGGTTGCAGTTGTCGCGCACGCCTCTGCCCGCCAGCCCGCTCCACCGCCGTAG  
CGCCCGAGTGTCGGGGGGCGCACCCGAGTCGGGGCC**ATG**AGGCCGGGAACCGCGCTACAGGCCG  
TGCTGCTGGCCGTGCTGCTGGTGGGGCTGCGGGCCGCGACGGGTGCGCTGCTGAGTGCCTCGG  
ATTTGGACCTCAGAGGAGGGCAGCCAGTCTGCCGGGGAGGGACACAGAGGCCTTGTTATAAAG  
TCATTTACTTCCATGATACTTCTCGAAGACTGAACTTTGAGGAAGCCAAAAGAAGCCTGCAGGA  
GGGATGGAGGCCAGCTAGTCAGCATCGAGTCTGAAGATGAACAGAAACTGATAGAAAAGTTCA  
TTGAAAACCTCTTGCCATCTGATGGTGACTTCTGGATTGGGCTCAGGAGGCGTGAGGAGAAAC  
AAAGCAATAGCACAGCCTGCCAGGACCTTTATGCTTGGACTGATGGCAGCATATCACAATTTA  
GGAAGTGGTATGTGGATGAGCCGTCCTGCGGCAGCGAGGTCTGCGTGGTCATGTACCATCAGC  
CATCGGCACCCCGCTGGCATCGGAGGCCCTACATGTTCCAGTGGAATGATGACCGGTGCAACA  
TGAAGAACAATTTTCAATTTGCAAATATTCTGATGAGAAACCAGCAGTTTCTTCTAGAGAAGCTG  
AAGGTGAGGAAACAGAGCTGACAACACCTGTACTTCCAGAAGAAACACAGGAAGAAGATGCCA  
AAAAAACATTTAAAGAAAGTAGAGAAGCTGCCTTGAATCTGGCCTACATCCTAATCCCCAGCA  
TTCCCTTCTCCTCCTCCTTGTGGTCACCACAGTTGTATGTTGGGTTTGGATCTGTAGAAAAA  
GAAAACGGGAGCAGCCAGACCTAGCACAAAGAAGCAACACACCATCTGGCCCTCTCCTCACC  
AGGGAAACAGCCCGGACCTAGAGGTCTACAATGTCATAAGAAAACAAAGCGAAGCTGACTTAG  
CTGAGACCCGGCCAGACCTGAAGAATATTTCAATCCGAGTGTGTTCCGGGAGAAGCCACTCCCG  
ATGACATGTCTTGTGACTATGACAACATGGCTGTGAACCCATCAGAAAGTGGGTTTGTGACTC  
TGGTGAGCGTGGAGAGTGGATTTGTGACCAATGACATTTATGAGTTCTCCCCAGACCAAATGG  
GGAGGAGTAAGGAGTCTGGATGGGTGGAAAATGAAATATATGGTTAT**TAG**GACATATAAAAAA  
CTGAAACTGACAACAATGGAAAAGAAATGATAAGCAAAATCCTCTTATTTTCTATAAGGAAAA  
TACACAGAAGGTCTATGAACAAGCTTAGATCAGGTCCTGTGGATGAGCATGTGGTCCCCACGA  
CCTCCTGTTGGACCCCCACGTTTTTGGCTGTATCCTTTATCCCAGCCAGTCATCCAGCTCGACC  
TTATGAGAAGGTACCTTGCCCAGGTCTGGCACATAGTAGAGTCTCAATAAATGTCACTTGGTT  
GGTTGTATCTAACTTTTAAGGGACAGAGCTTTACCTGGCAGTGATAAAGATGGGCTGTGGAGC  
TTGGAAAACACCTCTGTTTTTCTTGCTCTATACAGCAGCACATATTATCATAACAGACAGAAA  
ATCCAGAATCTTTTCAAAGCCCACATATGGTAGCACAGGTTGGCCTGTGCATCGGCAATTCTC  
ATATCTGTTTTTTTCAAAGAATAAAATCAAATAAAGAGCAGGAAAAAAAAA

20/615

**FIGURE 20**

MRPGTALQAVLLAVLLVGLRAATGRLLSASDLDLRGGQPVCRRGTQRPCYKVIYFHDTSRRLN  
FEEAKEACRRDGGQLVSI ESEDEQKLIKFIENLLPSDGDGFWIGLRRREEKQSNSTACQDLYA  
WTDGSISQFRNWWYVDEPSCGSEVCVVMYHQPSAPAGIGGPYMFQWNDDRCNMKNNFICKYSDE  
KPAVPSREAEGETELTTPVLPEETQEEDAKKTFKESREAAALNLAYILIPSIPLLLLLLVVTTV  
VCWWICRKRKREQPD PSTKKQHTIWPSPHQGN SPDLEVYNVIRKQSEADLAETRPDLKNISF  
RVCSGEATPDDMSCDYDNMAVNPSESGFVTLVSVESGFVTNDIYEFSPDQMGRSKESGWVENE  
IYGY

**Signal sequence:**

amino acids 1-21

**Transmembrane domain:**

amino acids 235-254

**N-glycosylation site.**

amino acids 117-121, 312-316

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 296-300

**Casein kinase II phosphorylation site.**amino acids 28-32, 30-34, 83-87, 100-104, 214-218, 222-226,  
299-303, 306-310, 323-327**N-myristoylation site.**

amino acids 18-24, 37-43, 76-82, 146-152

21/615

**FIGURE 21**

AGGCTCCCGCGCGCGGCTGAGTGCGGACTGGAGTGGGAACCCGGGTCCCGCGCTTAGAGAACACGCG**ATG**ACCA  
CGTGGAGCCTCCGGCGGAGGCCGGCCGACGCTGGGACTCCTGCTGCTGGTTCGTCTTGGGCTTCCTGGTGCTCC  
GCAGGCTGGACTGGAGCACCTTGGTCCCTCTGCGGCTCCGCCATCGACAGCTGGGGCTGCAGGCCAAGGGCTGGA  
ACTTCATGCTGGAGGATTCCACCTTCTGGATCTTCGGGGGCTCCATCCACTATTTCCGTGTGCCAGGGAGTACT  
GGAGGGACCGCCTGCTGAAGATGAAGGCCTGTGGCTTGAACACCCCTCACCACCTATGTTCCGTGGAACTGCATG  
AGCCAGAAAGAGGCAAATTTGACTTCTCTGGGAACCTGGACCTGGAGGCCTTCGTCTGATGGCCGCAGAGATCG  
GGCTGTGGGTGATTCTGCGTCCAGGCCCTACATCTGCAGTGAGATGGACCTCGGGGGCTTGCCCAGCTGGCTAC  
TCCAAGACCCCTGGCATGAGGCTGAGGACAACCTTACAAGGGCTTCACCGAAGCAGTGGACCTTTATTTTGACCACC  
TGATGTCCAGGGTGGTGCCACTCCAGTACAAGCGTGGGGGACCTATCATTGCCGTGCAGGTGGAGAATGAATATG  
GTTCTATAATAAAGACCCCGCATACATGCCCTACGTCAAGAAGGCACTGGAGGACCGTGGCATTGTGGAACCTGC  
TCCTGACTTCAGACAACAAGGATGGGCTGAGCAAGGGGATTGTCCAGGGAGTCTTGCCACCATCAACTTGCAGT  
CAACACACGAGCTGCAGCTACTGACCACCTTTCTCTTCAACGTCCAGGGGACTCAGCCCAAGATGGTGATGGAGT  
ACTGGACGGGGTGGTTTGAAGTCTGAGGAGGCCCTCACAATATCTTGGATTCTTCTGAGGTTTTGAAAACCGTGT  
CTGCCATTGTGGACGCCGGCTCCTCCATCAACCTCTACATGTTCCACGAGGCACCAACTTTGGCTTCATGAATG  
GAGCCATGCACTTCCATGACTACAAGTCAGATGTCAACAGCTATGACTATGATGCTGTGCTGACAGAAGCCGGCG  
ATTACACGGCCAAGTACATGAAGCTTCGAGACTTCTTCGGCTCCATCTCAGGCATCCCTCTCCCTCCCCACCTG  
ACCTTCTTCCCAAGATGCCGTATGAGCCCTTAACGCCAGTCTTGTAACCTGTCTCTGTGGGACGCCCTCAAGTACC  
TGGGGGAGCCAATCAAGTCTGAAAAGCCCATCAACATGGAGAACCTGCCAGTCAATGGGGGAAATGGACAGTCCCT  
TCGGGTACATTCTCTATGAGACCAGCATCACCTCGTCTGGCATCCTCAGTGGCCACGTGCATGATCGGGGGCAGG  
TGTTTGTGAACACAGTATCCATAGGATTCTTGGACTACAAGACAACGAAGATTGCTGTCCCCCTGATCCAGGGTT  
ACACCGTGCTGAGGATCTTGGTGGAGAATCGTGGGCGAGTCAACTATGGGGAGAATATTGATGACCAGCGCAAAG  
GCTTAATTGGAAATCTCTATCTGAATGATTACCCCTGAAAACTTCAGAATCTATAGCCTGGATATGAAGAAGA  
GCTTCTTTCAGAGGTTCCGCCTGGACAAATGGNGTTCCCTCCAGAAACACCCACATTACCTGCTTTCTTCTTGG  
GTAGCTTGTCCATCAGCTCCACGCCCTTGTGACACCTTTCTGAAGCTGGAGGGCTGGGAGAAGGGGGTTGTATTCA  
TCAATGGCCAGAACCTTGGACGTTACTGGAACATTGGACCCCAAGACGCTTTACCTCCCAGGTCCCTGGTTGA  
GCAGCGGAATCAACCAGGTCACTGTTTTTGGAGGACGATGGCGGGCCCTGCATTACAGTTCACGGAAACCCCC  
ACCTGGGCAGGAACCACTACATTAAG**TGAG**CGGTGGCACCCCTCCTGCTGGTGCCAGTGGGAGACTGCCGCCCTC  
CTCTTGACCTGAAGCCTGGTGGCTGCTGCCCCACCCCTCACTGCAAAAGCATCTCCTTAAGTAGCAACCTCAGGG  
ACTGGGGGCTACAGTCTGCCCCGTCTCAGCTCAAAACCCCTAAGCCTGCAGGGAAGGTGGGATGGCTCTGGGCC  
TGGCTTTGTGATGATGGCTTTCTTACAGCCCTGCTCTTGTGCCGAGGCTGTGCGGCTGTCTTAGGGTGGGAGC  
AGCTAATCAGATCGCCCAGCCTTTGGCCCTCAGAAAAAGTGCTGAAACGTGCCCTTGACCCGGACGTACAGCCC  
TGCGAGCATCTGCTGGACTCAGGCGTGCTCTTTGCTGGTTCTGGGAGGCTTGCCACATCCCTCATGGCCCCAT  
TTTATCCCCGAAATCCTGGGTGTGTACCAAGTGTAGAGGGTGGGGAAGGGGTGTCTCACCTGAGCTGACTTTGTT  
CTTCCTTCAACCTTCTGAGCCTTCTTTGGGATTCTGGAAGGAACTCGGCGTGAGAAACATGTGACTTCCCTTT  
TCCCTTCCCACTCGCTGCTTCCCACAGGGTGACAGGCTGGGCTGGAGAAACAGAAATCTCACCTGCGTCTTCC  
CAAGTTAGCAGGTGTCTCTGGTGTTCAGTGAGGAGGACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCA  
CATCCAGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCC  
AGGGAGGAGGACAGAAGGCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGA  
GGAGGACAGAAGGCCAGCTCACATGTGAGTCCTGGCAGAAGCCATGGCCCATGTCTGCACATCCAGGGAGGAGG  
ACAGAAGGCCAGCTCAGTGGCCCCGCTCCCCACCCCCACGCCGAACAGCAGGGGCAGAGCAGCCCTCCTTC  
GAAGTGTGTCCAAGTCCGCATTTGAGCCTTGTTCTGGGGCCAGCCCAACACCTGGCTTGGGCTCACTGTCTGA  
GTTGCAGTAAAGCTATAACCTTGAATCACA



22/615

**FIGURE 22**

MTTWSLRRRPPARTLGLLLLVVLGFLVLRRLDWSTLVPLRLRHRQLGLQAKGWNFMLEDSTFWI  
FGGSIHYFRVPREYWRDRLKMKACGLNTLTITYVPWNLHEPERGKFDGSGNLDLEAFVLMMAE  
IGLWVILRPGPYICSEMDLGGLP SWLLQDPGMRLRTTYKGFT EAVDLYFDHLSRVVPLQYKR  
GGPIIAVQVENEYGSYNKDPAYMPYVKKALEDRGIVELLTSDNKGDSLKGIVQGVLATINLQ  
STHELQLLTTFLEFNVQGTQPKMVMEYWTGWFD SWGGPHNILD SSEVLKTVSAIVDAGSSINLY  
MFHGGTNFGFMNGAMHFHDYKSDVTSYDYDAVLTEAGDY TAKYMKLRDFFGSI SGIPLP PPD  
LLPKMPYEPLTPVLYLSLWDALKYLGEPIKSEKPINMENLPVNGGNGQSFGYILYETSITSSG  
ILSGHVHDRGQVFVNTV SIGFLDYKTTKIAVPLIQGYTVLRILVENRGRVNYGENIDDQRKGL  
IGNLYLNDSP LKNFRIYS LDMKKSFFQRFGLDKWXS L PETPTLP AFFLGSL SISSTPCDTFLK  
LEGWEKGVVFINGQNLGRYWNIGPQKTL YLP GPWLSSGINQVIVFEETMAGPALQFTETPHLG  
RNQYIK

**Signal sequence:**

amino acids 1-27

**Casein kinase II phosphorylation site.**

amino acids 141-118, 253-257, 340-344, 395-399, 540-544, 560-564

**N-myristoylation site.**amino acids 146-152, 236-242, 240-246, 244-250, 287-293, 309-315,  
320-326, 366-372, 423-429, 425-431, 441-447, 503-509, 580-586

23 / 615

**FIGURE 23**

[illegible]

24/615

**FIGURE 24**

MNSSKSSETQCTERGCFSSQMFLWTVAGIPILFLSACFITRCVVTFRIFQTCDEKKFQLPENF  
TELSYNYGSGSVKNCCPLNWEYFQSSCYFFSTDTISWALSLKNCSAMGAHLVVINSQEEQEF  
LSYKKPKMREFFIGLSDQVVEGQWQWVDGTPLTKSLSFWDVGEPNNIATLEDCA TMRDSSNPR  
QNWNDVTCFLNYFRICEMVGINPLNKGKSL

**Signal sequence:**

amino acids 1-42

**N-glycosylation site.**

amino acids 2-6, 62-66, 107-111

**Casein kinase II phosphorylation site.**

amino acids 51-55, 120-124, 163-167, 175-179, 181-185

**N-myristoylation site.**

amino acids 15-21, 74-80, 155-161

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 27-38

25/615

**FIGURE 25**

GGGGACGCGGAGCTGAGAGGCTCCGGGCTAGCTAGGTGTAGGGGTGGACGGGTCCCAGGACCC  
TGGTGAGGGTTCTCTACTTGGCCTTCGGTGGGGGTCAAGACGCAGGCACCTACGCCAAAGGGG  
AGCAAAGCCGGGCTCGGCCCCGAGGCCCCCAGGACCTCCATCTCCCAATGTTGGAGGAATCCGA  
CACGTGACGGTCTGTCCGCGCTCTCAGACTAGAGGAGCGCTGTAAACGCCATGGCTCCCAAGA  
AGCTGTCTGCCTTCGTTCCCTGCTGCTGCCGCTCAGCCTGACGCTACTGCTGCCCCAGGCAG  
ACACTCGGTCGTTCTAGTGGATAGGGGTTCATGACCGGTTTCTCCTAGACGGGGCCCCGTTCC  
GCTATGTGTCTGGCAGCCTGCACTACTTTCCGGGTACCGCGGGTGCTTTGGGCCGACCGGCTTT  
TGAAGATGCGATGGAGCGGCCTCAACGCCATACAGTTTTATGTGCCCTGGAACCTACCACGAGC  
CACAGCCTGGGGTCTATAACTTTAATGGCAGCCGGGACCTCATTGCCTTTCTGAATGAGGCAG  
CTCTAGCGAACCTGTTGGTCATACTGAGACCAGGACCTTACATCTGTGCAGAGTGGGAGATGG  
GGGGTCTCCCATCCTGGTTGCTTCGAAAACCTGAAATTCATCTAAGAACCTCAGATCCAGACT  
TCCTTGCCGAGTGGACTCCTGGTTCAAGGTCTTGCTGCCCAAGATATATCCATGGCTTTATC  
ACAATGGGGGCAACATCATTAGCATTCAAGGTGGAGAATGAATATGGTAGCTACAGAGCCTGTG  
ACTTCAGCTACATGAGGCACTTGGCTGGGCTCTTCCGTGCACTGCTAGGAGAAAAGATCTTGC  
TCTTCACCACAGATGGGCCTGAAGGACTCAAGTGTGGCTCCCTCCGGGGACTCTATACCACTG  
TAGATTTTGGCCCAGCTGACAACATGACCAAATCTTTACCCTGCTTCGGAAGTATGAACCCC  
ATGGGCCATTGGTAAACTCTGAGTACTACACAGGCTGGCTGGATTACTGGGGCCAGAATCACT  
CCACACGGTCTGTGTCAGCTGTAACCAAAGGACTAGAGAACATGCTCAAGTTGGGAGCCAGTG  
TGAACATGTACATGTTCCATGGAGGTACCAACTTTGGATATTGGAATGGTGGCGATAAGAAGG  
GACGCTTCCTTCCGATTACTACCAGCTATGACTATGATGCACCTATATCTGAAGCAGGGGACC  
CCACACCTAAGCTTTTGTCTCTTCGAGATGTCATCAGCAAGTTCCAGGAAGTTCTTTGGGAC  
CTTTACCTCCCCCGAGCCCCAAGATGATGCTTGGACCTGTGACTCTGCACCTGGTTGGGCATT  
TACTGGCTTTCTAGACTTGCTTTGCCCCCGTGGGCCCATTCATTCAATCTTGCCAATGACCT  
TTGAGGCTGTCAAGCAGGACCATGGCTTCATGTTGTACCGAACCTATATGACCCATACCATTT  
TTGAGCCAACACCATTCTGGGTGCCAAATAATGGAGTCCATGACCGTGCCTATGTGATGGTGG  
ATGGGGTGTTCAGGGTGTGTTGGGAGCGAAATATGAGAGACAACTATTTTTGACGGGGAAAC  
TGGGGTCCAACTGGATATCTTGGTGGAGAACATGGGGAGGCTCAGCTTTGGGTCTAACAGCA  
GTGACTTCAAGGGCCTGTTGAAGCCACCAATTCTGGGGCAAACAATCCTTACCCAGTGGATGA  
TGTTCCCTCTGAAAATTGATAACCTTGTGAAGTGGTGGTTTCCCCTCCAGTTGCCAAAATGGC  
CATATCCTCAAGCTCCTTCTGGCCCCACATTCTACTCCAAAACATTTCCAATTTTAGGCTCAG  
TTGGGGACACATTTCTATATCTACCTGGATGGACCAAGGGCCAAGTCTGGATCAATGGGTTTA  
ACTTGGGCCGGTACTGGACAAAGCAGGGGCCACAACAGACCCTCTACGTGCCAAGATTCCTGC  
TGTTTCCTAGGGGAGCCCTCAACAAAATTACATTGCTGGAAGTAGAAGATGTACCTCTCCAGC  
CCCAAGTCCAATTTTGGATAAGCCTATCCTCAATAGCACTAGTACTTTGCACAGGACACATA  
TCAATTCCCTTTCAGCTGATACACTGAGTGCCTCTGAACCAATGGAGTTAAGTGGGCACTGAA  
AGGTAGGCCGGGCATGGTGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTGAGACGGGTG  
GATTACCTGAGGTCAGGACTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCGTCTCCACTA  
AAAATACAAAATTAGCCGGGCGTGATGGTGGGCACCTCTAATCCCAGCTACTTGGGAGGCTG  
AGGGCAGGAGAATTGCTTGAATCCAGGAGGCAGAGGTTGCAGTGAGTGGAGGTTGTACCACTG  
CACTCCAGCCTGGCTGACAGTGAGACACTCCATCTCAAAAAAAAAAAAA

26/615

**FIGURE 26**

MAPKKLSCLRSLLLPLSLTLLLPQADTRSFVVDGRGHDREFLLDGAPFRYVSGSLHYFRVPRVLW  
ADRLKMRWSGLNAIQFYVPWNYHEPQPGVYNFNGSRDLIAFLNEAALANLLVILRPGPYICA  
EWEMGGLPSWLLRKPEIHLRTSDPDFLAAVDSWFKVLLPKIYPWLYHNGGNIISIQVENEYGS  
YRACDFS YMRHLAGLFRALLGEKILLFTTDGPEGLKCGSLRGLYTTVDGFPADNMTKIFTLLR  
KYEPHGPLVNSEYYTGWLDYWGQNHSTRSVSAVTKGLENMLKLGASVNMVMFHGGTNFGYWNG  
ADKKGRFLPITTSYDYDAPISEAGDPTPKLFALRDVISKFQEVPLGPLPPSPKMMLGPVTLH  
LVGHLLAFLDLLCPRGPIHSILPMTFEAVKQDHGFMLYRTYMTHTIFEPTPFWVPNNGVHDRA  
YVMVDGVFQGVVERNMRDKLFLTGKLGSKLDILVENMGRLSFGSNSSDFKGLLKPPILGQIL  
TQWMMFPLKIDNLVKKWFPLQLPKWPYPQAPSGPTFYSKTFPILGSVGDTFLYLPGWTKGQVW  
INGFNLGRYWTQGPQOTLYVPRFLLFPRGALNKITLLELEDVPLQPQVQFLDKPILNSTSTL  
HRTHINSLSADTLASASEPMELSGH

**Signal sequence:**

amino acids 1-27

**N-glycosylation site.**

amino acids 97-101, 243-247, 276-280, 486-490, 625-629

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 4-8

**Casein kinase II phosphorylation site.**amino acids 148-152, 234-238, 327-331, 423-427, 469-473, 550-554,  
603-607, 644-648**Tyrosine kinase phosphorylation site.**

amino acids 191-198

**N-myristoylation site.**amino acids 131-137, 176-182, 188-194, 203-209, 223-229, 227-233,  
231-237, 274-280, 296-300, 307-313, 447-453, 484-490

27 / 615

## FIGURE 27

[illegible]

28/615

**FIGURE 28**

MGLLLLVPLLLLPGSYGLPFYNGFYYSNSANDQNLGNGHGKDLLNGVKLVVETPEETLFTYQG  
ASVILPCRYRYEPALVSPRRVRVKWWKLENGAPEKDVLVAIGLRHRSFGDYQGRVHLRQDKE  
HDVSLEIQDLRLLEDYGRYRCEVIDGLEDESGLVELELRGVVFPYQSPNGRYQFNHHEGQQVCA  
EQAAVVASFEQLFRAWEEGLDWCNAGWLQDATVQYPIMLPRQPCGGPGLAPGVRSYGPRHRL  
HRYDVFCFATALKGRVYYLEHPEKLTLEAREACQEDDATIAKVGQLFAAWKFHGLDRCDAGW  
LADGSVRYPPVHHPNCGPPEPGVRSFGFPDPQSRLYGVYCYRQH

**Signal sequence:**

amino acids 1-17

**Casein kinase II phosphorylation site.**

amino acids 29-33, 53-57, 111-115, 278-282

**Tyrosine kinase phosphorylation site.**

amino acids 137-145

**N-myristoylation site.**amino acids 36-42, 184-190, 208-214, 237-243, 297-303,  
307-313

29/615

**FIGURE 29**

GCAAGCGGCGAAATGCGGCCCTCCGGGAGTCTTGCAAGTTCCCCTGGCAGTCCTGGTGCTGTTG  
CTTTGGGGTGCTCCCTGGACGCACGGGCGGCGGAGCAACGTTTCGCGTCATCACGGACGAGAAC  
TGGAGAGAACTGCTGGAAGGAGACTGGATGATAGAATTTTATGCCCCGTGGTGCCCTGCTTGT  
CAAAATCTTCAACCGGAATGGGAAAGTTTTGCTGAATGGGGAGAAGATCTTGAGGTTAATATT  
GCGAAAGTAGATGTCACAGAGCAGCCAGGACTGAGTGGACGGTTTATCATAACTGCTCTTCCT  
ACTATTTATCATTGTAAAGATGGTGAATTTAGGCGCTATCAGGGTCCAAGGACTAAGAAGGAC  
TTCATAAACTTTATAAGTGATAAAGAGTGGAAGAGTATTGAGCCCGTTTCATCATGGTTTGGT  
CCAGGTTCTGTTCTGATGAGTAGTATGTCAGCACTCTTTCAGCTATCTATGTGGATCAGGACG  
TGCCATAACTACTTTATTGAAGACCTTGGATTGCCAGTGTGGGGATCATATACTGTTTTTGTCT  
TTAGCAACTCTGTTTTCCGGACTGTTATTAGGACTCTGTATGATATTTGTGGCAGATTGCCTT  
TGTCCTTCAAAAAGGCGCAGACCACAGCCATACCCATACCCTTCAAAAAAATTATTATCAGAA  
TCTGCACAACCTTTGAAAAAAGTGGAGGAGGAACAAGAGGCGGATGAAGAAGATGTTTCAGAA  
GAAGAAGCTGAAAGTAAAGAAGGAACAAACAAAGACTTTCCACAGAATGCCATAAGACAACGC  
TCTCTGGGTCCATCATTGGCCACAGATAAATCCTAGTTAAATTTTATAGTTATCTTAATATTA  
TGATTTTGATAAAACAGAAGATTGATCATTTTGTGGTTTGAAGTGAAGTGTGACTTTTTT  
GAATATTGCAGGGTTCAGTCTAGATTGTCAATTAATTGAAGAGTCTACATTCAGAACATAAAA  
GCACTAGGTATACAAGTTTGAAATATGATTTAAGCACAGTATGATGGTTTAAATAGTTCTCTA  
ATTTTTGAAAAATCGTGCCAAGCAATAAGATTTATGTATATTTGTTTAAATAATAACCTATTTC  
AAGTCTGAGTTTTGAAAATTTACATTTCCCAAGTATTGCATTATTGAGGTATTTAAGAAGATT  
ATTTTAGAGAAAAATATTTCTCATTTGATATAATTTTTCTCTGTTTCACTGTGTGAAAAAAG  
AAGATATTTCCCATAAATGGGAAGTTTGCCATTGTCTCAAGAAATGTGTATTTCACTGACAA  
TTTCGTGGTCTTTTTAGAGGTATATTCCAAAATTTCCCTTGATTTTTTAGGTTATGCAACTAAT  
AAAACTACCTTACATTAATTAATTACAGTTTTCTACACATGGTAATACAGGATATGCTACTG  
ATTTAGGAAGTTTTTAAGTTCATGGTATTCTCTTGATTCCAACAAAGTTTGATTTTCTCTTGT  
ATTTTTCTTACTTACTATGGGTACATTTTTTATTTTTCAAATTGGATGATAATTTCTTGGAA  
ACATTTTTTATGTTTTAGTAAACAGTATTTTTTTGTTGTTTCAAAGTGAAGTTTACTGAGAGA  
TCCATCAAATTGAACAATCTGTTGTAATTTAAATTTTGGCCACTTTTTTCAGATTTTACATC  
ATTCCTTGCTGAACCTCAACTGAAATTGTTTTTTTTTTCTTTTTGGATGTGAAGGTGAACATT  
CCTGATTTTTGTCTGATGTGAAAAAGCCTTGGTATTTTACATTTTGAAAAATTCAAAGAAGCTT  
AATATAAAAGTTTGCATTCTACTCAGGAAAAAGCATCTTCTGTATATGTCTTAAATGTATTT  
TTGTCTCATATACAGAAAGTTCTTAATTGATTTTACAGTCTGTAATGCTTGATGTTTTAAAA  
TAATAACATTTTTATATTTTTTAAAGACAACTTCATATTATCCTGTGTCTTTCTGACTG  
GTAATATTGTGTGGGATTTTACAGGTAAAAGTCAGTAGGATGGAACATTTTAGTGTATTTTA  
CTCCTTAAAGAGCTAGAATACATAGTTTTTACCTTAAAGAAGGGGGAAAAATCATAAATACAA  
TGAATCAACTGACCATACGTAGTAGACAATTTCTGTAATGTCCCCTTCTTCTAGGCTCTGT  
TGCTGTGTGAATCCATTAGATTTACAGTATCGTAATATACAAGTTTTCTTAAAGCCCTCTCC  
TTTAGAATTTAAATATTGTACCATTAAAGAGTTTGGATGTGTAAGTGTGATGCCTTAGAAA  
AATATCCTAAGCACAAAATAAACCTTTCTAACCACCTCATTAAAGCTGAAAAA



30/615

**FIGURE 30**

MAPSGSLAVPLAVLVLLLWGAPWTHGRRSNVRVITDENWRELLEGDWMIEFYAPWCPACQNLO  
PEWESFAEWGEDLEVNIKVDVTEQPGLSGRFIITALPTIYHCKDGEFRRYQGPRTKKDFINF  
ISDKEWKSIEPVSSWFGPGSVLMSSMSALFQLSMWIRTCHNYFIEDLGLPVWGSYTVFALATL  
FSGLLLGLCMIFVADCLCPSKRRRPQYPYPYPSKKLLSESAQPLKKVEEEQEADVEDVSEEEAE  
SKEGTNKDFPQNAIRQRS LGPSLATDKS

**Signal sequence:**

amino acids 1-26

**Transmembrane domain:**

amino acids 182-201

**Casein kinase II phosphorylation site.**

amino acids 68-72, 119-123, 128-132, 247-251, 257-261

**Tyrosine kinase phosphorylation site.**

amino acids 107-115

**N-myristoylation site.**

amino acids 20-26, 192-198

**Amidation site.**

amino acids 25-29

31/615

**FIGURE 31**

AG**ATG**GCGGTCTTGGCACCTCTAATTGCTCTCGTGTATTTCGGTGCCGCGACTTTTCACGATGGC  
TCGCCCCAACCTTACTACCTTCTGTGCGGCCCTGCTCTCTGCTGCCTTCCTACTCGTGAGGAAAC  
TGCCGCGCTCTGCCACGGTCTGCCACCCAACGCGAAGACGGTAACCCGTGTGACTTTGACT  
GGAGAGAAGTGGAGATCCTGATGTTTTCTCAGTGCCATTGTGATGATGAAGAACCGCAGATCCA  
TCACTGTGGAGCAACATATAGGCAACATTTTCATGTTTAGTAAAGTGGCCAACACAATTCTTT  
TCTTCCGCTTGGATATTTCGCATGGGCCTACTTTACATCACACTCTGCATAGTGTTCTTGATGA  
CGTGCAAACCCCCCTATATATGGGCCCTGAGTATATCAAGTACTTCAATGATAAAACCATTG  
ATGAGGAAGTAGAACGGGACAAGAGGGTCACCTGGATTGTGGAGTTCTTTGCCAATTGGTCTA  
ATGACTGCCAATCATTTGCCCTATCTATGCTGACCTCTCCCTTAAATACAAGTGTACAGGGC  
TAAATTTTGGGAAGGTGGATGTTGGACGCTATACTGATGTTAGTACGCGGTACAAAGTGAGCA  
CATCACCCCTCACCAAGCAACTCCCTACCCTGATCCTGTTCCAAGGTGGCAAGGAGGCAATGC  
GGCGGCCACAGATTGACAAGAAAGGACGGGCTGTCTCATGGACCTTCTCTGAGGAGAATGTGA  
TCCGAGAATTTAACTTAAATGAGCTATAACCAGCGGGCCAAGAACTATCAAAGGCTGGAGACA  
ATATCCCTGAGGAGCAGCCTGTGGCTTCAACCCCCACCACAGTGTGAGATGGGGAAAACAAGA  
AGGATAAA**TAA**GATCCTCACTTTGGCAGTGCTTCCTCTCCTGTCAATTCCAGGCTCTTTCCAT  
AACCACAAGCCTGAGGCTGCAGCCTTTNATTNATGTTTTCCCTTTGGCTGNGACTGGNTGGGG  
CAGCATGCAGCTTCTGATTTTAAAGAGGCATCTAGGGAATTGTCAGGCACCCTACAGGAAGGC  
CTGCCATGCTGTGGCCAACTGTTTCACTGGAGCAAGAAAGAGATCTCATAGGACGGAGGGGGA  
AATGGTTTTCCCTCCAAGCTTGGGTCAAGTGTGTTAACTGCTTATCAGCTATTCAGACATCTCCA  
TGGTTTCTCCATGAACTCTGTGGTTTCATCATTCCCTCTTAGTTGACCTGCACAGCTTGGTT  
AGACCTAGATTTAACCCTAAGGTAAGATGCTGGGGTATAGAACGCTAAGAATTTTCCCCAAG  
GACTCTTGCTTCCTTAAGCCCTTCTGGCTTCGTTTATGGTCTTCATTAAAAGTATAAGCCTAA  
CTTTGTGCTAGTCCTAAGGAGAAACCTTTAACCACAAAGTTTTATCATTGAAGACAATATT  
GAACAACCCCTATTTTGTGGGGATTGAGAAGGGGTGAATAGAGGCTTGAGACTTTCTTTGT  
GTGGTAGGACTTGGAGGAGAAATCCCCTGGACTTTCACCTAACCCCTCTGACATACTCCCCACAC  
CCAGTTGATGGCTTTCCGTAATAAAAAGATTGGGATTCCTTTTG

32/615

**FIGURE 32**

MAVLAPLIALVYSVPRLSRWLAQPYYLLSALLSAAFLLVRLPPLCHGLPTQREDGNPCDFDW  
REVEILMFLSAIVMMKNRRSITVEQHIGHNIFMFSKVANTILFFRLDIRMGLLYITLCIVFLMT  
CKPPLYMGPEYIKYFNDKTIDEELERDKRVTWIVEFFANWSNDCQSFAPIYADLSLKYNCTGL  
NFGKVDVGRYTDVSTRYKVSTSP LTKQLPTLILFQGGKEAMRRPQIDKKGRAVSWTFSEENVI  
REFNLNELYQRAKKLSKAGDNIPEEQPVASTPTTVSDGENKKDK

**Signal sequence:**

amino acids 1-48

**Transmembrane domain:**

amino acids 111-125

**N-glycosylation site.**

amino acids 165-169, 185-189

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 154-158, 265-269

**Casein kinase II phosphorylation site.**

amino acids 51-55, 145-149, 245-249, 286-290, 288-292

**N-myristoylation site.**

amino acids 188-194, 225-231

**Myb DNA-binding domain repeat signature 1.**

amino acids 244-253

33/615

**FIGURE 33**

CGGACGCGTGGGGTGCCCGAC**ATG**GCGAGTGTAGTGCTGCCGAGCGGATCCCAGTGTGCGGCG  
GCAGCGGCGGCGGCGGCGCCTCCCGGGCTCCGGCTTCTGCTGTTGCTCTTCTCCGCCGCGGCA  
CTGATCCCCACAGGTGATGGGCAGAATCTGTTTACGAAAGACGTGACAGTGATCGAGGGAGAG  
GTTGCGACCATCAGTTGCCAAGTCAATAAGAGTGACGACTCTGTGATTGAGTACTGAATCCC  
AACAGGCAGACCATTATTTTCAGGGACTTCAGGCCTTTGAAGGACAGCAGGTTTCAGTTGCTG  
AATTTTTCTAGCAGTGAACCTCAAAGTATCATTGACAAACGTCTCAATTTCTGATGAAGGAAGA  
TACTTTTGCCAGCTCTATAACCGATCCCCACAGGAAAGTTACACCACCATCACAGTCCTGGTC  
CCACCACGTAATCTGATGATCGATATCCAGAAAGACACTGCGGTGGAAGGTGAGGAGATTGAA  
GTCAACTGCACTGCTATGGCCAGCAAGCCAGCCACGACTATCAGGTGGTTCAAAGGGAACACA  
GAGCTAAAAGGCAAATCGGAGGTGGAAGAGTGGTCAGACATGTACACTGTGACCAGTCAGCTG  
ATGCTGAAGGTGCACAAGGAGGACGATGGGGTCCCAGTGATCTGCCAGGTGGAGCACCCCTGCG  
GTCAGTGGAAACCTGCAGACCCAGCGGTATCTAGAAGTACAGTATAAGCCTCAAGTGCACATT  
CAGATGACTTATCCTCTACAAGGCTTAACCCGGGAAGGGGACGCGCTTGAGTTAACATGTGAA  
GCCATCGGGAAGCCCCAGCCTGTGATGGTAACCTTGGGTGAGAGTCGATGATGAAATGCCTCAA  
CACGCCGTACTGTCTGGGCCCAACCTGTTTCATCAATAACCTAAACAAAACAGATAATGGTACA  
TACCGCTGTGAAGCTTCAAACATAGTGGGGAAAGCTCACTCGGATTATATGCTGTATGTATAC  
GATCCCCCACAACCTATCCCTCCTCCCACAACAACCACCACCACCACCACCACCACCACCACC  
ACCATCCTTACCATCATCACAGATTCCTCGAGCAGGTGAAGAAGGCTCGATCAGGGCAGTGGAT  
CATGCCGTGATCGGTGGCGTCGTGGCGGTGGTGGTGTTCGCCATGCTGTGCTTGCTCATCATT  
CTGGGGCGCTATTTTGCCAGACATAAAGGTACATACTTCACTCATGAAGCCAAAGGAGCCGAT  
GACGCAGCAGACGCAGACACAGCTATAATCAATGCAGAAGGAGGACAGAACAACCTCCGAAGAA  
AAGAAAGAGTACTTCATCT**AGAT**ATCAGCCTTTTTGTTTCAATGAGGTGTCCAACCTGGCCCTATT  
TAGATGATAAAGAGACAGTGATATTGG

34/615

**FIGURE 34**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39518
<subunit 1 of 1, 440 aa, 1 stop
<MW: 48240, pI: 4.93, NX(S/T): 7
MASVVLPSGSQCAAAAAAAPPGLRLLLLLFSAAALIPTGDGQNLFTKDVTVIEGEVATISCQ
VNKSDDSVIQLLNPNRQTIYFRDFRPLKDSRFQLLNFSSSELKVSLTNVSISDEGRYFCQLYT
DPPQESYTTITVLVPPRNLMDIQKDTAVEGEEIEVNCTAMASKPATTIRWFKGNTELKGKSE
VEEWSDMYTVTSQMLMKVHKEDDGVPVICQVEHPAVTGNLQTQRYLEVQYKPQVHIQMTYPLQ
GLTREGDALELTCEAIGKPQPMVTWVRVDDEMPQHAVLSGPNLFINNLNKTDNGTYRCEASN
IVGKAHSDYMLYVYDPPTTIPPTTTTTTTTTTTTTTILTIITDSRAGEEGSIRAVDHAIVIGGV
VAVVVFAMLCLLIILGRYFARHKGTYFTHEAKGADDAADATAINAEGGQNNSEEKKEYFI
```

**Signal sequence.**

amino acids 1-36

**Transmembrane domain.**

amino acids 372-393

**N-glycosylation sites.**amino acids 65-69, 99-103, 111-115, 163-167, 302-306, 306-310,  
430-434**Tyrosine kinase phosphorylation sites.**

amino acids 233-240, 319-328

**N-myristoylation sites.**amino acids 9-15, 227-233, 307-313, 365-371, 376-382, 402-408,  
411-417, 427-433, 428-432

35/615

**FIGURE 35**

GGTTGCCACAGCTGGTTTtagggcccccGACCACTGGGGCCOCTTGTcaggaggagacAGCCTCCCGGCCCGGGGAG  
GACAAGTCGCTGCCACCTTTGGCTGCCGACGTGATTCCTTGGGACGGTCCGTTTCCTGCCGTcAGCTGCCGGCCG  
AGTTGGGTCTCCGTGTTTCAGGCCGGCTCCCCCTTCTGGTCTCCCTTCTCCCGCTGGGCCGGTTTATCGGGAGG  
AGATTGTCTTCCAGGGCTAGCAATTGGACTTTTGATGATGTTTGACCCAGCGGCAGGAATAGCAGGCAACGTGAT  
TTCAAAGCTGGGCTCAGCCTCTGTTTCTTCTCTCGTGTAATCGCAAAACCCATTTTGAGCAGGAATTCCAATCA  
**TGT**CTGTGATGGTGGTGAGAAAGAAGGTGACACGGAAATGGGAGAACTCCAGGCAGGAACACCTTTTGCTGTG  
ATGGCCGCGTCATGATGGCCCGCAAAAGGGCATTCTTCTACCTGACCCTTTTCCTCATCCTGGGGACATGTACAC  
TCTTCTTCGCCCTTTGAGTGCCGCTACCTGGCTGTTcAGCTGTCTCCTGCCATCCCTGTATTTGCTGCCATGCTCT  
TCCTTTTCTCCATGGCTACACTGTTGAGGACCAGCTTcAGTGACCCTGGAGTGATTCTCGGGCGCTACCAGATG  
AAGCAGCTTTCATAGAAATGGAGATAGAAGCTACCAATGGTGCGGTGCCCCAGGGCCAGCGACCACCGCTCGTA  
TCAAGAATTTCCAGATAAAACAACCAGATTGTGAACTGAAATACTGTTACACATGCAAGATCTTCGGCCTCCCC  
GGGCCTCCCATTGCAGCATCTGTGACAACTGTGTGGAGCGCTTCGACCATCACTGCCCCCTGGGTGGGGAATTGTG  
TTGGAAAGAGGAACTACCGCTACTTCTACCTCTTCATCCTTTCTCTCTCCCTCCTCACAATCTATGTCTTCGCCCT  
TCAACATCGTCTATGTGGCCCTCAAATCTTTGAAAATTGGCTTCTTGAGACATTGAAAGAACTCCTGGAACTG  
TTCTAGAAGTCCTCATTGCTTCTTTACACTCTGGTCCGTCGTGGGACTGACTGGATTTCTACTTTCTCGTGG  
CTCTCAACCAGACAACCAATGAAGACATCAAAGGATCATGGACAGGGAAGAATCGCGTCCAGAATCCCTACAGCC  
ATGGCAATATTGTGAAGAACTGCTGTGAAGTGCTGTGTGGCCOCTTGCCCCCAGTGTGCTGGATCGAAGGGGTA  
TTTTGCCACTGGAGGAAAGTGGAACTCGACCTCCAGTACTCAAGAGACCAGTAGCAGCCTCTTGCCACAGAGCC  
CAGCCCCACAGAACACCTGAACTCAAATGAGATGCCGGAGGACAGCAGCACTCCCGAAGAGATGCCACCTCCAG  
AGCCCCCAGAGCCACCACAGGAGGCAGCTGAAGCTGAGAAG**TAG**CCTATCTATGGAAGAGACTTTTGTGTGTG  
TAATTAGGGCTATGAGAGATTTcAGGTGAGAAGTTAAACCTGAGACAGAGAGCAAGTAAGCTGTCCCTTTTAACT  
GTTTTTCTTTGGTCTTTAGTCACCAGTTGCACACTGGCATTCTTCTGCTGCAAGCTTTTTTAAATTTCTGAACT  
CAAGGCAGTGGCAGAAGATGTCAGTCACCTCTGATAACTGGAAAAATGGGTCTCTTGGGCCCTGGCACTGGTTCT  
CCATGGCCTCAGCCACAGGGTCCOCTTGGAACCCCTCTCTTCCCTCCAGATCCAGCCCTCCTGCTTGGGGTCAC  
TGGTCTCATTCTGGGGCTAAAAGTTTTTGAGACTGGCTCAAATCCTCCCAAGCTGCTGCACGTGCTGAGTCCAGA  
GGCAGTCACAGAGACCTCTGGCCAGGGGATCCTAACTGGGTCTTGGGGTCTTCAGGACTGAAGAGGAGGGAGAG  
TGGGGTCAGAAGATTCTCCTGGCCACCAAGTGCCAGCATTGCCACAAATCCTTTTAGGAATGGGACAGGTACCT  
TCCACTTGTTGTANNNNNNNNNNNNNNNNNNNNNNNNTGTTTTCTTTTGACTCCTGCTCCCATTAGGAG  
CAGGAATGGCAGTAATAAAAGTCTGCACTTTGGTCATTTCTTTTCTCAGAGGAAGCCCGAGTGCTCACTTAAAC  
ACTATCCCCCTCAGACTCCCTGTGTGAGGCCTGCAGAGGCCCTGAATGCACAAATGGGAAACCAAGGCACAGAGAG  
GCTCTCCTCTCCTCTCCTCTCCCCGATGTACCCTCAAAAAAAAAAAAAATGCTAACCAGTTCTTCCATTAAAGCCT  
CGGCTGAGTGAGGGAAAGCCCAGCACTGCTGCCCTCTCGGGTAACTCACCTAAGGCCTCGGCCACCTCTGGCT  
ATGGTAACCACACTGGGGGCTTCTCCAAGCCCCGCTCTTCCAGCACTTCCACCGGCAGAGTCCAGAGCCACTT  
CACCTGGGGGTGGGCTGTGGCCCCAGTCAGCTCTGCTCAGGACCTGCTCTATTTcAGGGAAGAAGATTTATGT  
ATTATATGTGGCTATATTTCTAGAGCACCTGTGTTTTCTCTTTCTAAGCCAGGGTCTGTCTGGATGACTTAT  
CGGGTGGGGGAGTGTAACCGGAACTTTTCTATCTATTTGAAGGCGATTAACTGTGTCTAATGCA

36/615

**FIGURE 36**

MSVMVVRKKVTRKWEKLPGRNTFCCDGRVMMARQKGIFYLTLFLILGTCTLFFAFECRYLAVQ  
LSPAIPVFAAMLFLFSMATLLRTSFSDPGVIPRALPDEAAFIEMEIEATNGAVPQGGQRPPPRI  
KNFQINNQIVKLKYCYTCKIFRPPRASHCSICDNCVERFDHHC PWVGNCVGKRNRYFYLFIL  
SLSLLTIYVFAFNIVYVALKSLKIGFLETLKETPGTVLEVLI CFFTLWSVGLTGFHTFLVAL  
NQT T NEDIKGSWTGKNRVQNPYSHGNIVKNCCEVLCGPLPPSVLDRRGILPLEESGSRPPSTQ  
ETSSSLLPQSPAPTEHLNSNEMPEDSSTPEEMP PPEPPEPPQEAAEAEK

**Putative transmembrane domains:**

amino acids 36-55 (type II TM), 65-84, 188-208, 229-245

37/615

**FIGURE 37**

GGCGGAGCAGCCCTAGCCGCCACCGTCGCTCTCGCAGCTCTCGTCGCCACTGCCACCGCCGCCGCCGTCCTACTGCG  
TCCTGGCTCCGGCTCCCGCGCCCTCCCGGCCGGCCATGCGAGCCCCGCCGCGCCAGGCGCCCGGTGCGCAGCTGC  
TGCCCCGCGCTGGCCCTGCTGCTGCTGCTGCTCGGAGCGGGGCCCGAGGCAGCTCCCTGGCCAAACCGGTGCCCG  
CCGCGCCCTTGTCTGCGCCCGGGCCGTGCGCCGCGCAGCCCTGCCGGAATGGGGGTGTGTGCACCTCGCGCCCTG  
AGCCGGACCCGAGCACCCGGCCCCCGCGCGGAGCCTGGCTACAGCTGCACCTGCCCCGCCGGGATCTCCGGCG  
CCAACTGCCAGCTTGTGAGATCCTTGTGCCAGCAACCCCTTGTCAACATGGCAACTGCAGCAGCAGCAGCAGCA  
GCAGCAGCGATGGCTACCTCTGCATTTGCAATGAAGGCTATGAAGGTCCCAACTGTGAACAGGCACTTCCAGTC  
TCCAGCCACTGGCTGGACCGAATCCATGGCACCCCGACAGCTTACGCTGTTCTGCTACTCAGGAGCCTGACA  
AAATCCTGCCTCGCTCTCAGGCAACGGTGACACTGCCTACCTGGCAGCCGAAAACAGGGCAGAAAGTTGTAGAAA  
TGAAATGGGATCAAGTGGAGGTGATCCCAGATATTGCCTGTGGGAATGCCAGTTCTAACAGCTCTGCGGGTGGCC  
GCCTGGTATCCTTTGAAGTGCCACAGAACACCTCAGTCAAGATTGCGCAAGATGCCACTGCCTCACTGATTTTGC  
TCTGGAAGGTCACGGCCACAGGATTCCAACAGTGCTCCCTCATAGATGGACGAAGTGTGACCCCTTTCAGGCTT  
CAGGGGGACTGGTCCTCCTGGAGGAGATGCTCGCCTTGGGGAATAATCACTTTATTGGTTTTGTGAATGATTCTG  
TGACTAAGTCTATTGTGGCTTTGCGCTTAACCTCTGGTGGTGAAGGTCAGCACCTGTGTGCCGGGGGAGAGTCACG  
CAAATGACTTGGAGTGTTCAGGAAAAGGAAAATGCACCACGAAGCCGTGAGGGCAACTTTTTCTGTACCTGTG  
AGGAGCAGTACGTGGGTACTTTCTGTGAAGAATACGATGCTTGCCAGAGGAAACCTTGCCAAAACAAACGCGAGCT  
GTATTGATGCAATGAAAAGCAAGATGGGAGCAATTTACCTGTGTTTGCCTTCTGGTTATACTGGAGAGCTTT  
GCCAGTCCAAGATTGATTACTGCATCCTAGACCCATGCAGAAATGGAGCAACATGCATTTCCAGTCTCAGTGGAT  
TCACCTGCCAGTGTCCAGAAGGATACTTCGGATCTGCTTGTGAAGAAAAGGTGGACCCCTGCGCCTCGTCTCCGT  
GCCAGAACAACGGCACCTGCTATGTGGACGGGGTACACTTTACCTGCAACTGCAGCCCGGGCTTCACAGGGCCGA  
CCTGTGCCAGCTTATTGACTTCTGTGCCCTCAGCCCTGTGCTCATGGCACGTGCCGACGCTGGGCACCAGCT  
ACAAATGCCTCTGTGATCCAGGTTACCATGGCCTCTACTGTGAGGAGGAATATAATGAGTGCCTCTCCGCTCCAT  
GCCTGAATGCAGCCACCTGCAGGGACCTCGTTAATGGCTATGAGTGTGTGTGCCTGGCAGAATACAAAGGAACAC  
ACTGTGAATTGTACAAGGATCCCTGCGCTAACGTCACTGTCTGAACGGAGCCACCTGTGACAGCGACGGCCTGA  
ATGGCACGTGCATCTGTGCACCCGGGTTTACAGGTGAAGAGTGCGACATTGACATAAATGAATGTGACAGTAACC  
CCTGCCACCATGGTGGGAGCTGCCTGGACCAGCCCAATGGTTATAACTGCCACTGCCCGCATGGTTGGGTGGGAG  
CAAATGTGAGATCCACCTCCAATGGAAGTCCGGGCACATGGCGGAGAGCCTCACCAACATGCCACGGCACTCCC  
TCTACATCATCATTGGAGCCCTCTGCGTGGCCTTCATCCTTATGCTGATCATCCTGATCGTGGGGATTTGCCGCA  
TCAGCCGCATTGAATACCAGGGTTCTTCCAGGCCAGCCTATGAGGAGTTCTACAACCTGCCGACGATCGACAGCG  
AGTTTCAGCAATGCCATTGCATCCATCCGGCATGCCAGGTTTGGAAAGAAATCCCGGCCCTGCAATGTATGATGTA  
GCCCCATCGCCTATGAAGATTACAGTCCCTGATGACAAACCCTTGGTCACACTGATTAAGAACTAAAGATTTGTAAAT  
CTTTTTTTGGATTATTTTTCAAAAAGATGAGATACTACACTCATTTAAATATTTTTAAGAAAATAAAAAGCTTAA  
GAAATTTAAATGCTAGCTGCTCAAGAGTTTTTCAGTAGAATATTTAAGAACTAATTTTCTGCAGCTTTTAGTTTG  
GAAAAAATATTTTTAAAAACAAATTTGTGAAACCTATAGACGATGTTTTAATGTACCTTCAGCTCTCTAAACTGT  
GTGCTTCTACTAGTGTGTGCTCTTTTCACTGTAGACACTATCACGAGACCCAGATTAATTTCTGTGGTTGTTACA  
GAATAAGTCTAATCAAGGAGAAGTTTCTGTTTGACGTTTGAGTGCCGGCTTCTGAGTAGAGTTAGGAAAACCAC  
GTAACTAGCATATGATGTATAATAGAGTATACCCGTTACTTTAAAAAGAAAGTCTGAAATGTTTCGTTTTGTGGAAA  
AGAACTAGTTAAATTTACTATTCTTAACCCGAATGAAATTAGCCTTTGCCTTATTCTGTGCATGGGTAAGTAAC  
TTATTTCTGCAGTGTGTTGTGAACTTTGTGGAACATTCTTTCGAGTTTGTGTTTGTGCTATTTCGTAACAGTCG  
TCGAACTAGGCCTCAAAAACATACGTAACGAAAAGGCCCTAGCGAGGCAAATTCGATTGATTTGAATCTATATTT  
TTCTTTAAAAAGTCAAGGGTTCTATATTGTGAGTAAATTAATTTACATTTGAGTTGTTTGTGCTAAGAGGTAG  
TAAATGTAAGAGAGTACTGGTTCCTTCAGTAGTGAGTATTTCTCATAGTGCAGCTTTATTTATCTCCAGGATGTT  
TTTGTGGCTGTATTGATTGATATGTGCTTCTTCTGATTCTTGCTAATTTCCAACCATATTGAATAAATGTGATC  
AAGTCA



38/615

**FIGURE 38**

><subunit 1 of 1, 737 aa, 1 stop  
 ><MW: 78475, pI: 5.09, NX(S/T): 11  
 MQPRRAQAPGAQLLPALALLLLLLGAGPRGSSLANPVPAAPLSAPGPCAAQPCRNGGVCTSRP  
 EPDPQHPAPAGEPGYSCTCPAGISGANCQLVADPCASNPCHHGNCSSSSSSSSSDGYLCICNEG  
 YEGPNCEQALPSLPATGWTESMAPRQLQVPATQEPDKILPRSQATVTLPTWQPKTGQKVEM  
 KWDQVEVIPDIACGNASSNSSAGGRLVSFEVPQNTSVKIRQDATASLILLWKVTATGFQQCSL  
 IDGRSVTPLQASGGLVLEEMLALGNNHFIGFVNDVSVTKSIVALRLTLVVKVSTCVPGESHAN  
 DLECSGKGKCTTKPSEATFSCTCEEQYVGTFCEEYDACQRKPCQNNASCIDANEKQDGSNFTC  
 VCLPGYTGELCQSKIDYCIDLPCRNGATCISSLSGFTCQCPEGYFGSACEEKVDPCASSPCQN  
 NGTCYVDGVHFTCNCSPGFTGPTCAQLIDFCALSPCAHGTCSRVTGTSYKCLCDPGYHGLYCEE  
 EYNECLSAPCLNAATCRDLVNGYECVCLA EYKGTHCELYKDPCANVSCLNGATCDS DGLNGTC  
 ICAPGFTGEECDIDINECDSNPCHHGGSCLDQPNGYNCHCPHGWVGANCEIHLQWKS GHMAES  
 LTNMPRHSLYIIIGALCVAFILMLIILIVGICRISRIEYQGSSRPAYEEFYNCRSIDSEFSNA  
 IASIRHARFGKSRPAMYDVSP IAYEDYSPDDKPLVTLIKTKDL

**Signal sequence.**

amino acids 1-28

**Transmembrane domain.**

amino acids 641-660

**N-glycosylation sites.**amino acids 107-111, 204-208, 208-212, 223-227, 286-290, 361-365,  
375-379, 442-446, 549-553, 564-568**Glycosaminoglycan attachment site.**

amino acids 320-324

**Tyrosine kinase phosphorylation sites.**

amino acids 490-498, 674-682

**N-myristoylation sites.**amino acids 30-36, 56-62, 57-63, 85-91, 106-112, 203-209,  
373-379, 449-455, 480-486, 562-568, 565-571**Amidation site.**

amino acids 702-706

**Aspartic acid and asparagine hydroxylation site.**

amino acids 520-532, 596-608

**EGF-like domain cysteine pattern signatures.**amino acids 80-92, 121-133, 336-348, 378-390, 416-428, 454-466,  
491-503, 529-541, 567-579, 605-617

39/615

**FIGURE 39**

[illegible]

40/615

**FIGURE 40**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA39979

&gt;&lt;subunit 1 of 1, 204 aa, 1 stop

&gt;&lt;MW: 22147, pI: 8.37, NX(S/T): 3

MVCGGFACSKNCLCALNLLYTLVSLLLIGIAAWGIGFGLISSLRVVGVIAGVIFLFLIALVG  
LIGAVKHHQVLLFFYMIILLVFIIVQFSVSCACLALNQEQQGQLLEVGNNTASARNDIQRN  
NCCGFRSVNPNDTCLASCVKSDHSCSPCAPIIGEYAGEVLRVGGIGLFFSFTEILGVWLT  
YRNQKDPRANPSAFL

**Signal Peptide:**

amino acids 1-34

**Transmembrane domains:**

amino acids 47-63, 72-95 and 162-182

41/615

**FIGURE 41**

CAGTCACCA**ATGA**AGCTGGGCTGTGTCCTCATGGCCTGGGCCCTCTACCTTTCCCTTGGTGTGC  
TCTGGGTGGCCAGATGCTACTGGCTGCCAGTTTTTGAGACGCTGCAGTGTGAGGGACCTGTCT  
GCACTGAGGAGAGCAGCTGCCACACGGAGGATGACTTGACTGATGCAAGGGAAGCTGGCTTCC  
AGGTCAAGGCCTACACTTTTCAGTGAACCCCTCCACCTGATTGTGTCCTATGACTGGCTGATCC  
TCCAAGGTCCAGCCAAGCCAGTTTTTTGAAGGGGACCTGCTGGTTCTGCGCTGCCAGGCCTGGC  
AAGACTGGCCACTGACTCAGGTGACCTTCTACCGAGATGGCTCAGCTCTGGGTCCCCCGGGC  
CTAACAGGGAATTCTCCATCACCGTGGTACAAAAGGCAGACAGCGGGCACTACCACTGCAGTG  
GCATCTTCCAGAGCCCTGGTCTGGGATCCCAGAAACAGCATCTGTTGTGGCTATCACAGTCC  
AAGAAGTGTTCAGCGCCAATTCTCAGAGCTGTACCCTCAGCTGAACCCCAAGCAGGAAGCC  
CCATGACCCTGAGTTGTTCAGACAAAGTTGCCCTGCAGAGGTGAGCTGCCCGCCTCCTCTTCT  
CCTTCTACAAGGATGGAAGGATAGTGCAAAGCAGGGGGCTCTCCTCAGAATTCCAGATCCCCA  
CAGCTTCAGAAGATCACTCCGGGTCATACTGGTGTGAGGCAGCCACTGAGGACAACCAAGTTT  
GGAAACAGAGCCCCCAGCTAGAGATCAGAGTGCAGGGTGCTTCCAGCTCTGCTGCACCTCCCA  
CATTGAATCCAGCTCCTCAGAAATCAGCTGCTCCAGGAAGTCTCCTGAGGAGGCCCTGGGC  
CTCTGCCTCCGCCGCCAACCCCATCTTCTGAGGATCCAGGCTTTTCTTCTCCTCTGGGGATGC  
CAGATCCTCATCTGTATCACAGATGGGCCTTCTTCTCAAACACATGCAGGATGTGAGAGTCC  
TCCTCGGTACCTGCTCATGGAGTTGAGGGAATTATCTGGCCACCAGAAGCCTGGGACCACAA  
AGGCTACTGCTGA**ATGA**AGTAAACAGTTCATCCATGATCTCACTTAACCACCCCAATAAATC  
TGATTCTTTATTTTCTCTTCTGTCCTGCACATATGCATAAGTACTTTTACAAGTTGTCCAG  
TGTTTTGTTAGAATAATGTAGTTAGGTGAGTGTAAATAAATTTATATAAAGTGAGAATTAGAG  
TTTAGCTATAATTGTGTATTCTCTCTTAACACAACAGAATTCTGCTGTCTAGATCAGGAATTT  
CTATCTGTTATATCGACCAGAATGTTGTGATTTAAAGAGAACTAATGGAAGTGGATTGAATAC  
AGCAGTCTCAACTGGGGGCAATTTTGGCCCCCAGAGGACATTGGGCAATGTTTGGAGACATTT  
TGGTCATTATACTTGGGGGGTTGGGGGATGGTGGGATGTGTGTCTACTGGCATCCAGTAAATA  
GAAGCCAGGGGTGCCGCTAAACATCCTATAATGCACAGGGCAGTACCCACAACGAAAAATAA  
TCTGGCCCCAAATGTCAGTTGTACTGAGTTTGAGAAACCCAGCCTAATGAAACCCTAGGTGT  
TGGGCTCTGGAATGGGACTTTGTCCCTTCTAATTATTATCTCTTTCCAGCCTCATTAGCTAT  
TCTTACTGACATAACAGTCTTTAGCTGGTGCTATGGTCTGTTCTTTAGTTCTAGTTTGTATCC  
CCTCAAAAGCCATTATGTTGAAATCCTAATCCCCAAGGTGATGGCATTAGAAGTGGGCCTTT  
GGGAAGTGATTAGATCAGGAGTGCAGAGCCCTCATGATTAGGATTAGTGCCCTTATTTAAAAA  
GGCCCCAGAGAGCTAACTCACCCCTCCACCATATGAGGACGTGGCAAGAAGATGACATGTATG  
AGAACCAAAAAACAGCTGTCGCCAAACACCGACTCTGTCGTTGCCTTGATCTTGAAGTTCCAG  
CCTCCAGAACTATGAGAAATAAAATCTGGTTGTTTGTAGCCTAA

42/615

**FIGURE 42**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA40594
><subunit 1 of 1, 359 aa, 1 stop
><MW: 38899, pI: 5.21, NX(S/T): 0
MKLGCVLMAWALYLSLGLWVAQMLLAASFETLQCEGPVCTEESSCHTEDDLTDAREAGFQVK
AYTFSEPFHLIVSYDWLILQGPAKPVFEGDLLVLRQAWQDWPLTQVTFYRDGSALGPPGPNR
EFSITVVQKADSGHYHCSGIFQSPGPGIPETASVVAITVQELFPAPILRAVPSAEPQAGSPMT
LSCQTKLPLQRSAARLLFSFYKDGRIVQSRGLSSEFQIPTASEDHSGSYWCEAATEDNQVWKQ
SPQLEIRVQGASSAAPTLLNPAPQKSAAPGTAPEEAPGPLPPPPTPSSDPGFSSPLGMPDP
HLYHQMGLLLKHMQDVRVLLGHLLMELRELSGHQKPGTTKATAE
```

**Signal sequence:**

amino acids 1-17

**Leucine zipper pattern sequence:**

amino acids 12-33

**Protein kinase C phosphorylation site:**

amino acids 353-355

43/615

**FIGURE 43**

GCGAGTGTCCAGCTGCGGAGACCCGTGATAATTCGTAACTAATTCAACAAACGGGACCCTTC  
TGTGTGCCAGAAACCGCAAGCAGTTGCTAACCCAGTGGGACAGGCGGATTGGAAGAGCGGGAA  
GGTCCTGGCCCAGAGCAGTGTGACACTTCCCTCTGTGACCATGAACTCTGGGTGTCTGCATT  
GCTGATGGCCTGGTTTGGTGTCTGAGCTGTGTGCAGGCCGAATTCTTCACCTCTATTGGGCA  
CATGACTGACCTGATTTATGCAGAGAAAGAGCTGGTGCAGTCTCTGAAAGAGTACATCCTTGT  
GGAGGAAGCCAAGCTTTCCAAGATTAAGAGCTGGGCCAACAAAATGGAAGCCTTGACTAGCAA  
GTCAGCTGCTGATGCTGAGGGCTACCTGGCTCACCTGTGAATGCCTACAACTGGTGAAGCG  
GCTAAACACAGACTGGCCTGCGCTGGAGGACCTTGTCTGCAGGACTCAGCTGCAGGTTTTAT  
CGCCAACCTCTCTGTGCAGCGGCAGTTCTTCCCCACTGATGAGGACGAGATAGGAGCTGCCAA  
AGCCCTGATGAGACTTCAGGACACATACAGGCTGGACCCAGGCACAATTTCCAGAGGGGAAC  
TCCAGGAACCAAGTACCAGGCAATGCTGAGTGTGGATGACTGCTTTGGGATGGGCCGCTCGGC  
CTACAATGAAGGGGACTATTATCATACGGTGTGTGGATGGAGCAGGTGCTAAAGCAGCTTGA  
TGCCGGGGAGGAGGCCACCACAACCAAGTCACAGGTGCTGGACTACCTCAGCTATGCTGTCTT  
CCAGTTGGGTGATCTGCACCGTGCCCTGGAGCTCACCCGCCGCTGCTCTCCCTTGACCCAAG  
CCACGAACGAGCTGGAGGGAATCTGCGGTACTTTGAGCAGTTATTGGAGGAAGAGAGAGAAAA  
AACGTTAACAAATCAGACAGAAGCTGAGCTAGCAACCCAGAAAGGCATCTATGAGAGGCCTGT  
GGACTACCTGCCTGAGAGGGATGTTTACGAGAGCCTCTGTCTGGGGAGGGTGTCAAACCTGAC  
ACCCCGTAGACAGAAGAGGCTTTTCTGTAGGTACCACCATGGCAACAGGGCCCCACAGCTGCT  
CATTGCCCCCTTCAAAGAGGAGGACGAGTGGGACAGCCCGCACATCGTCAGGTACTACGATGT  
CATGTCTGATGAGGAAATCGAGAGGATCAAGGAGATCGCAAAACCTAAACTTGCACGAGCCAC  
CGTTCGTGATCCCAAGACAGGAGTCTCACTGTGCCAGCTACCGGGTTTCCAAAAGCTCCTG  
GCTAGAGGAAGATGATGACCCTGTTGTGGCCCGAGTAAATCGTCGGATGCAGCATATCACAGG  
GTTAACAGTAAAGACTGCAGAATTGTTACAGGTTGCAAATTATGGAGTGGGAGGACAGTATGA  
ACCGCACTTCGACTTCTCTAGGCGACCTTTTGACAGCGGCCTCAAACAGAGGGGAATAGGTT  
AGCGACGTTTCTTAACCTACATGAGTGATGTAGAAGCTGGTGGTGCCACCGTCTTCCCTGATCT  
GGGGGCTGCAATTTGGCCTAAGAAGGGTACAGCTGTGTTCTGGTACAACCTCTTGCGGAGCGG  
GGAAGGTGACTACCGAACAAGACATGCTGCCTGCCCTGTGCTTGTGGGCTGCAAGTGGGTCTC  
CAATAAGTGGTTCCATGAACGAGGACAGGAGTTCTTGAGACCTTGTGGATCAACAGAAGTTGA  
CTGACATCCTTTTCTGTCCTTCCCCTTCCCTGGTCCTTCAGCCCATGTCAACGTGACAGACACC  
TTTGTATGTTCTTTGTATGTTCTTATCAGGCTGATTTTTGGAGAAATGAATGTTTGTCTGGA  
GCAGAGGGAGACCATACTAGGGCGACTCCTGTGTGACTGAAGTCCAGCCCTTCCATTACGCC  
TGTGCCATCCCTGGCCCCAAGGCTAGGATCAAAGTGGCTGCAGCAGAGTTAGCTGTCTAGCGC  
CTAGCAAGGTGCCTTTGTACCTCAGGTGTTTTAGGTGTGAGATGTTTCAGTGAACCAAAGTTC  
TGATACCTTGTTTACATGTTTGTATTTTATGGCATTCTATCTATTGTGGCTTTACCAAAAAAT  
AAAATGTCCCTACCAGAAAAA

44/615

**FIGURE 44**

MKLWVSALLMAWFGVLSCVQAEFFTSIGHMTDLIYAEEKELVQSLKEYIILVEEAKLSKIKSWAN  
KMEALTSKSAADAEGYLAHPVNAYKLVKRLNTDWPALDVLQDSAAGFIANLSVQRQFFPTD  
EDEIGAALKMRLQDITYRLDPGTISRGELPGTKYQAMLSVDDCFGMGRSAYNEGDIYHTVLWM  
EQVLKQLDAGEEATTTKSQVLDYLSYAVFQLGDLHRALELTRRLSLDPSHERAGGNLRYFEQ  
LLEEEREKTLTNQTEAELATPEGIYERPDVYELPERDVYESLCRGEVGLTPRRQKRLFCRYHH  
GNRAPQLLIAPFKEEWDSPHIVRYDVMSDEEIERIKEIAKPKLARATVRDPKTGVLTVAS  
YRVSKSSWLEEDDDPVVARVNRMQHITGLTVKTAELLQVANYGVGGQYEPHFDFSRRPFDSG  
LKTEGNRLATFLNYMSDVEAGGATVFPDLGAAIWPKKGTAVFWYNLLRSGEVDYRTRHAACPV  
LVGCKWVSNKWFHERGQEFRLPCGSTVD

**Signal sequence:**

amino acids 1-17

**N-glycosylation site.**

amino acids 115-119, 264-268

**Glycosaminoglycan attachment site.**

amino acids 490-494

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 477-481

**Casein kinase II phosphorylation site.**amino acids 43-47, 72-76, 125-129, 151-155, 165-169, 266-270,  
346-350, 365-369, 385-389, 457-461, 530-534**Tyrosine kinase phosphorylation site.**

amino acids 71-80, 489-496

**N-myristoylation site.**

amino acids 14-20, 131-137, 171-177, 446-452

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 8-19

**Leucine zipper pattern.**

amino acids 213-235

45/615

**FIGURE 45**

GGGGCCTTGCCCTCCGCACTCGGGCGCAGCCGGGTGGATCTCGAGCAGGTGCGGAGCCCCGGG  
CGGCGGGCGCGGGTGCGAGGGATCCCTGACGCCTCTGTCCCTGTTTCTTTGTGCTCCAGCC  
TGTCTGTCGTGTTTTGGCGCCCCCGCCTCCCCGCGGTGCGGGGTTGCACACCGATCCTGGGC  
TTCGCTCGATTTGCCGCCGAGGCGCCTCCCAGACCTAGAGGGGCGCTGGCCTGGAGCAGCGGG  
TCGTCTGTGTCCTCTCTCCTCTGCGCCGCGCCGGGGATCCGAAGGGTGCGGGGCTCTGAGGA  
GGTGACGCGCGGGGCTCCCGCACCCCTGGCCTTGCCCGCATCTCCCTCTCTCCAGGTGTGA  
GCAGCCTATCAGTCACC**ATG**TCCGCAGCCTGGATCCCGGCTCTCGGCCTCGGTGTGTGTCTGC  
TGCTGCTGCCGGGGCCCCGCGGGCAGCGAGGGAGCCGCTCCCATTGCTATCACATGTTTTACCA  
GAGGCTTGGACATCAGGAAAGAGAAAGCAGATGTCCTCTGCCAGGGGGCTGCCCTCTTGAGG  
AATTCTCTGTGTATGGGAACATAGTATATGCTTCTGTATCGAGCATATGTGGGGCTGCTGTCC  
ACAGGGGAGTAATCAGCAACTCAGGGGGACCTGTACGAGTCTATAGCCTACCTGGTCGAGAAA  
ACTATTCCTCAGTAGATGCCAATGGCATCCAGTCTCAAATGCTTTCTAGATGGTCTGCTTCTT  
TCACAGTAACTAAAGGCAAAAGTAGTACACAGGAGGCCACAGGACAAGCAGTGTCCACAGCAC  
ATCCACCAACAGGTAAACGACTAAAGAAAACACCCGAGAAGAAAACCTGGCAATAAAGATTGTA  
AAGCAGACATTGCATTTCTGATTGATGGAAGCTTTAATATTGGGCAGCGCCGATTTAATTTAC  
AGAAGAATTTTGTGGAAAAGTGGCTCTAATGTTGGGAATTGGAACAGAAGGACCACATGTGG  
GCCTTGTTCAAGCCAGTGAACATCCCAAAATAGAATTTTACTTGAAAACTTTACATCAGCCA  
AAGATGTTTTGTTGCCATAAAGGAAGTAGGTTTCAGAGGGGGTAATTCCAATACAGGAAAAG  
CCTTGAAGCATACTGCTCAGAAATTCTTCACGGTAGATGCTGGAGTAAGAAAAGGGATCCCCA  
AAGTGGTGGTGGTATTTATTGATGGTTGGCCTTCTGATGACATCGAGGAAGCAGGCATTGTGG  
CCAGAGAGTTTGGTGTCAATGTATTTATAGTTTCTGTGGCCAAGCCTATCCCTGAAGAACTGG  
GGATGGTTCAGGATGTCACATTTGTTGACAAGGCTGTCTGTCGGAATAATGGCTTCTTCTCTT  
ACCACATGCCCAACTGGTTTGGCACCAAAAATACGTAAAGCCTCTGGTACAGAAGCTGTGCA  
CTCATGAACAAATGATGTGCAGCAAGACCTGTTATAACTCAGTGAACATTGCCTTTCTAATTG  
ATGGCTCCAGCAGTGTTGGAGATAGCAATTTCCGCCTCATGCTTGAATTTGTTTCCAACATAG  
CCAAGACTTTTGAAATCTCGGACATTGGTGCCAAGATAGCTGCTGTACAGTTTACTTATGATC  
AGCGCACGGAGTTCAGTTTCACTGACTATAGCACCAAGAGAATGTCCTAGCTGTCATCAGAA  
ACATCCGCTATATGAGTGGTGGAACAGCTACTGGTGATGCCATTTCTTCACTGTTAGAAATG  
TGTTTGGCCCTATAAGGGAGAGCCCCAACAAGAACTTCCTAGTAATTGTCACAGATGGGCAGT  
CCTATGATGATGTCCAAGGCCCTGCAGCTGCTGCACATGATGCAGGAATCACTATCTTCTCTG  
TTGGTGTGGCTTGGGCACCTCTGGATGACCTGAAAGATATGGCTTCTAAACCGAAGGAGTCTC  
ACGCTTTCTTCAAGAGAGTTCACAGGATTAGAACCAATTGTTTCTGATGTCATCAGAGGCA  
TTTGTAGAGATTTCTTAGAATCCAGCAA**TAA**TGGTAACATTTTGACAACCTGAAAGAAAAAGT  
ACAAGGGGATCCAGTGTGTAATTTGTATTCTCATAATACTGAAATGCTTTAGCATACTAGAAT  
CAGATACAAAACCTATTAAGTATGTCAACAGCCATTTAGGCAAATAAGCACTCCTTTAAAGCCG  
CTGCCTTCTGGTTACAATTTACAGTGTACTTTGTTAAAAACACTGCTGAGGCTTCATAATCAT  
GGCTCTTAGAAACTCAGGAAAGAGGAGATAATGTGGATTAAACCTTAAGAGTTCTAACCATG  
CCTACTAAATGTACAGATATGCAAATTCATAGCTCAATAAAAGAATCTGATACTTAGACCAA  
AAAAA



46/615

**FIGURE 46**

MSAAWIPALGLGVCLLLLPGPAGSEGAAPIAITCFTRGLDIRKEKADVLCPGGCPLEEF SVYG  
NIVYASVSSICGAAVHRGVISNSGGPVRVYSLPGRENYSSVDANGIQSQMLSRWSASFTVTKG  
KSSTQEATGQAVSTAHPPTGKRLKKTPEKKTGNKDCKADIAFLIDGSFNIGQRRFNLQKNFVG  
KVALMLGIGTEGPHVGLVQASEHPKIEFYLNFTSAKDVLFAlKEVGFRGGNSNTGKALKHTA  
QKFFTVDAGVRKGIPKVVVVFIDGWPSDDIEEAGIVAREFGVNVFIVSVAKPIPEELGMVQDV  
TFVDKAVCRNNGFFSYHMPNWFGTTKYVKPLVQKLCTHEQMMCSKTCYNSVNIAFLIDGSSSV  
GDSNFRMLLEFVSNIAKTFEISDIGAKIAAVQFTYDQRTFSFTDYSTKENVLAVIRNIRYMS  
GGTATGDAISFTVRNVFGPIRESPNKNFLVIVTDGQSYDDVQGPAAAAHDAGITIFSVGVAVA  
PLDDLKDMASKPKESHAFFTREFTGLEPIVSDVIRGICRDFLESQQ

**Signal sequence:**

amino acids 1-24

**N-glycosylation site.**

amino acids 100-104, 221-225

**Casein kinase II phosphorylation site.**amino acids 102-106, 129-133, 224-228, 316-320, 377-381, 420-424,  
425-429, 478-482, 528-532**N-myristoylation site.**amino acids 10-16, 23-29, 81-87, 135-141, 158-164, 205-211,  
239-245, 240-246, 261-267, 403-409, 442-448, 443-449**Amidation site.**

amino acids 145-149

47/615

**FIGURE 47**

GCCCCGCGCCCGGCGCCGGGCGCCCCGAAGCCGGGAGCCACCGCC**ATG**GGGGCCTGCCTGGGAG  
CCTGCTCCCTGCTCAGCTGCGCGTCCTGCCTCTGCGGCTCTGCCCCCTGCATCCTGTGCAGCT  
GCTGCCCCGCCAGCCGCAACTCCACCGTGAGCCGCTCATCTTCACGTTCTTCCTCTTCCTGG  
GGGTGCTGGTGTCCATCATTTATGCTGAGCCCGGGCGTGGAGAGTCAGCTCTACAAGCTGCCCT  
GGGTGTGTGAGGAGGGGGCCGGGATCCCCACCGTCCTGCAGGGCCACATCGACTGTGGCTCCC  
TGCTTGGCTACCGCGCTGTCTACCGCATGTGCTTCGCCACGGCGGCCTTCTTCTTCTTCTTT  
TCACCTTGCTCATGCTCTGCGTGAGCAGCAGCCGGGACCCCCGGGCTGCCATCCAGAATGGGT  
TTTGGTTCTTTAAGTTCTTGATCCTGGTGGGCCTCACCGTGGGTGCCTTCTACATCCCTGACG  
GCTCCTTCACCAACATCTGGTTCTACTTCGGCGTCGTGGGCTCCTTCTTCTTCATCCTCATCC  
AGCTGGTGCTGCTCATCGACTTTGCGCACTCCTGGAACCAGCGGTGGCTGGGCAAGGCCGAGG  
AGTGCGATTCCCGTGCTGGTACGCAGGCCTCTTCTTCTTCACTCTCCTCTTCTACTTGCTGT  
CGATCGCGGCCGTGGCGCTGATGTTTATGTACTACACTGAGCCCAGCGGCTGCCACGAGGGCA  
AGGTCTTCATCAGCCTCAACCTCACCTTCTGTGTCTGCGTGTCCATCGCTGCTGTCTGCCCCA  
AGGTCCAGGACGCCCAGCCCAACTCGGGTCTGCTGCAGGCCTCGGTCATCACCTCTACACCA  
TGTTTGTACCTGGTCAGCCCTATCCAGTATCCCTGAACAGAAATGCAACCCCCATTTGCCAA  
CCCAGCTGGGCAACGAGACAGTTGTGGCAGGCCCCGAGGGCTATGAGACCCAGTGGTGGGATG  
CCCCGAGCATTGTGGGCCTCATCATCTTCTCCTGTGCACCCTCTTCATCAGTCTGCGCTCCT  
CAGACCACCGGCAGGTGAACAGCCTGATGCAGACCGAGGAGTGCCACCTATGCTAGACGCCA  
CACAGCAGCAGCAGCAGCAGGTGGCAGCCTGTGAGGGCCGGGCCTTTGACAACGAGCAGGACG  
GCGTCACCTACAGCTACTCCTTCTTCCACTTCTGCCTGGTGCTGGCCTCACTGCACGTCATGA  
TGACGCTCACCAACTGGTACAAGCCCGGTGAGACCCGGAAGATGATCAGCACGTGGACCGCCG  
TGTGGGTGAAGATCTGTGCCAGCTGGGCAGGGCTGCTCCTCTACCTGTGGACCCTGGTAGCCC  
CACTCCTCCTGCGCAACCGCGACTTCAGCT**TGAG**GGCAGCCTCACAGCCTGCCATCTGGTGCCTC  
CTGCCACCTGGTGCCTCTCGGCTCGGTGACAGCCAACCTGCCCCCTCCCCACACCAATCAGCC  
AGGCTGAGCCCCCACCCTGCCCCAGCTCCAGGACCTGCCCCCTGAGCCGGGCCTTCTAGTCGT  
AGTGCCCTCAGGGTCCGAGGAGCATCAGGCTCCTGCAGAGCCCCATCCCCCGCCACACCCAC  
ACGGTGGAGCTGCCTCTTCTTCCCCTCCTCCCTGTTGCCATACTCAGCATCTCGGATGAAA  
GGGCTCCCTTGTCTCAGGCTCCACGGGAGCGGGGCTGCTGGAGAGAGCGGGGAACTCCCACC  
ACAGTGGGGCATCCGGCACTGAAGCCCTGGTGTTCCTGGTCACGTCCCCCAGGGGACCCTGCC  
CCCTTCTGGACTTCGTGCCTTACTGAGTCTCTAAGACTTTTCTAATAAACAAGCCAGTGCG  
TGTAATAAAAAA

48/615

**FIGURE 48**

MGACLGACSLLSASCCLCGSAPCILCSCCPASRNSTVSRLIFTFFLFLGVLVSIIMLSPGVES  
QLYKLPWVCEEGAGIPTVLQGHIDCGSLLGYRAVYRMCFATAAFFFFFFFFTLLMLCVSSSRDPR  
AAIQNGFWFFKFLILVGLTVGAFYIPDGSTNIWFYFGVVGSFLFILIQLVLLIDFAHSWNQR  
WLGKAEECDSRAWYAGLFFFTLLFYLLSIAAVALMFMYYTEPSGCHEGKVFISLNLTFVCVVS  
IAAVLPKVQDAQPNISGLLQASVITLYTMFVTWSALSSIPEQKCNPHLPTQLGNETVVAGPEGY  
ETQWWDAPSIVGLIIFLLCTLFISLRSSDHRQVNSLMQTEECPPMLDATQQQQQQVAACEGRA  
FDNEQDGVTSYSFFHFCLVLASLHVMMTLTNWYKPGETRKMISTWTAVWVKICASWAGLLLY  
LWTLVAPLLLRNRDFS

**Signal sequence:**  
amino acids 1-20

**Transmembrane domains:**  
amino acids 40-58, 101-116, 134-150, 162-178, 206-223, 240-257,  
272-283, 324-340, 391-406, 428-444

49/615

**FIGURE 49**

[illegible]

50/615

**FIGURE 50**

MAGIPGLLFLFFLLCAVGQVSPYSAPWKPTWPAYRLPVVLPQSTLNLAKPDFGAEAKLEVSS  
SCGPQCHKGTPLPTYEEAKQYLSYETLYANGSRTETQVGIIYLSSSGDGAQHRDSGSSGKSRR  
KRQIYGYSRFSIFGKDFLLNYPFSTSVKLSTGCTGTLVAEKHVLTAAHCIHDGKTYVKGTQK  
LRVGFLKPKFKDGGRGANDSTSAMPEQMKFQWIRVKRTHVPGWIKGNANDIGMDYDYALLEL  
KKPHKRKFMKIGVSPPAKQLPGGRIHFSGYDNDRPGNLVYRFCDVKDETYDLLYQQCDAQPGA  
SGSGVYVRMWKRQQQKWERKIIGIFSGHQWDMNGSPQDFNVAVRITPLKYAQICYWIKGNYL  
DCREG

**Signal sequence:**

amino acids 1-19

**N-glycosylation site.**

amino acids 93-97, 207-211

**Glycosaminoglycan attachment site.**

amino acids 109-113, 316-320

**Casein kinase II phosphorylation site.**

amino acids 77-81, 95-99, 108-112, 280-284, 351-355

**N-myristoylation site.**

amino acids 159-165, 162-168, 202-208, 205-211, 314-320, 338-344

**Serine proteases, trypsin family, histidine active site.**

amino acids 171-177

51/615

**FIGURE 51**

GGGAGGGGGCTCCGGGCGCCGCGCAGCAGACCTGCTCCGGCCGCGCGCCTCGCCGCTGTCTCCGGGAGCGGCAG  
CAGTAGCCCCGGGCGGCGAGGGCTGGGGGTTCTCGAGACTCTCAGAGGGGCGCCTCCCATCGGCGCCCCACCACCC  
CAACCTGTTCTCGCGCGCCACTGCGCTGCGCCCCAGGACCCGCTGCCAACATGGATTTTCTCTGGCGCTGGT  
GCTGGTATCCTCGCTCTACCTGCAGGCGGCGCCGAGTTCGACGGGAGGTGGCCAGGCAATAGTGTATCGAT  
TGGCCTATGTGTTATGGTGGGAGGATTGACTGCTGCTGGGGCTGGGCTCGCCAGTCTTGGGGACAGTGTAGCC  
TGTGTGCCAACCACGATGCAAACATGGTGAATGTATCGGGCCAAACAAGTGCAAGTGTATCCTGGTTATGCTGG  
AAAAACCTGTAATCAAGATCTAAATGAGTGTGGCCTGAAGCCCCGGCCCTGTAAGCACAGGTGCATGAACACTTA  
CGGCAGCTACAAGTGCTACTGTCTCAACGGATATATGCTCATGCCGGATGGTTCTGCTCAAGTGCCCTGACCTG  
CTCCATGGCAAACCTGTCAGTATGGCTGTGATGTTGTTAAAGGACAAATACGGTGCCAGTGCCCATCCCCTGGCCT  
GCACCTGGCTCCTGATGGGAGGACCTGTGTAGATGTTGATGAATGTGCTACAGGAAGAGCCTCCTGCCCTAGATT  
TAGGCAATGTGTCAACACTTTTGGGAGCTACATCTGCAAGTGTCTATAAAGGCTTCGATCTCATGTATATTGGAGG  
CAAATATCAATGTCATGACATAGACGAATGCTCACTTGGTCAGTATCAGTGCAGCAGCTTTGCTCGATGTTATAA  
CGTACGTGGGTCCCTACAAGTGCAAATGTAAAGAAGGATACCAGGGTGATGGACTGACTTGTGTGTATATCCCAAA  
AGTTATGATTGAACCTTCAGGTCCAATTCATGTACCAAAGGGAATGGTACCATTTTAAAGGGTGACACAGGAAA  
TAATAATTGGATTCTGATGTTGGAAGTACTTGGTGGCTCCGAAGACACCATATATCTCTCTATCATTACCAA  
CAGGCCTACTTCTAAGCCAACAACAAGACCTACACCAAAGCCAACACCAATTCCTACTCCACCACCACCACCACC  
CCTGCCAACAGAGCTCAGAACACCTCTACCACCTACAACCCAGAAAGGCCAACACCACCGGACTGACAACTATAGC  
ACCAGCTGCCAGTACACCTCCAGGAGGGATTACAGTTGACAACAGGGTACAGACAGACCCTCAGAAACCCAGAGG  
AGATGTGTTTCAAGTGTCTGGTACACAGTTGTAATTTTGACCATGGACTTTGTGGATGGATCAGGGAGAAAGACAA  
TGACTTGCACTGGGAACCAATCAGGGACCCAGCAGGTGGACAATATCTGACAGTGTGGCAGCCAAAGCCCCAGG  
GGGAAAAGCTGCACGCTTGGTGCTACCTCTCGGCCGCTCATGCATTGAGGGGACCTGTGCTGTCTATTGAGGCA  
CAAGGTGACGGGGCTGCACTCTGGCACACTCCAGGTGTTTGTGAGAAAACACGGTGCCACGGAGCAGCCCTGTG  
GGGAAGAAATGGTGGCCATGGCTGGAGGCAAACACAGATCACCTTGCGAGGGGCTGACATCAAGAGCGAATCACA  
AAGATGATTAAAGGGTTGGAAAAAAGATCTATGATGGAAAATTAAGGAACTGGGATTATTGAGCCTGGAGAAG  
AGAAGACTGAGGGGCAAACCATTTGATGGTTTTCAAGTATATGAAGGGTTGGCACAGAGAGGGTGGCGACCAGCTG  
TTCTCCATATGCACTAAGAATAGAACAAGAGGAACTGGCTTAGACTAGAGTATAAGGGAGCATTCTTGGCAGG  
GGCCATTGTTAGAACTTCTATAAAAAAAGAGTGTGAAAATCTCAGTATCTCTCTCTCTTTCTAAAAAATTAGA  
TAAAAATTTGTCTATTTAAGATGGTTAAAGATGTTCTTACCCAAGGAAAAGTAACAAATTATAGAATTTCCCAAA  
AGATGTTTTGATCCTACTAGTAGTATGCAGTGAAAATCTTTAGAACTAAATAATTTGGACAAGGCTTAATTTAGG  
CATTTCCCTCTTGACCTCCTAATGGAGAGGGATTGAAAGGGGAAGAGCCCAACCAATGCTGAGCTCACTGAAATA  
TCTCTCCCTTATGGCAATCCTAGCAGTATTAAAGAAAAAGGAACTATTTATTCCAAATGAGAGTATGATGGAC  
AGATATTTTAGTATCTCAGTAATGTCCTAGTGTGGCGGTGGTTTTCAATGTTTCTTCATGGTAAAGGTATAAGCC  
TTTCATTTGTTCAATGGATGATGTTTCAGATTTTTTTTTTTTTTAAGAGATCCTTCAAGGAACACAGTTCAGAGAG  
ATTTTCATCGGGTGCAATCTCTCTGCTTCGTGTGTGACAAGTTATCTTGGCTGCTGAGAAAGAGTGCCCTGCCCC  
ACACCGGCAGACCTTTCTTCACTCATCAGTATGATTGAGTTTCTCTTATCAATTGGACTCTCCAGGTTCCAC  
AGAACAGTAATATTTTTTGAACAATAGGTACAATAGAAGGTCTTCTGTCAATTAACCTGGTAAAGGCAGGGCTGG  
AGGGGGAAAATAAATCATTAAGCCTTTGAGTAACGGCAGAAATATATGGCTGTAGATCCATTTTAAATGGTTCATT  
TCCTTTATGGTCATATAACTGCACAGCTGAAGATGAAAGGGGAAAATAAATGAAAATTTTACTTTTCGATGCCAA  
TGATACATTGCACTAAACTGATGGAAGAAGTTATCCAAAGTACTGTATAACATCTTGTGTTATTATTTAATGTTTT  
CTAAAATAAAAAATGTTAGTGGTTTTCCAAATGGCCTAATAAAAAACAATTATTTGTAAATAAAAAACTGTTAGTAAT

52/615

**FIGURE 52**

MDFLALVLVSSLYLQAAAEFDGRWPRQIVSSIGLCRYGGRIDCCWGWARQSWGQCQPVCQPR  
CKHGEICIGPNKCKCHPGYAGKTCNQDLNECGLKPRPCKHRCMNTYGSYKCYCLNGYMLMPDGS  
CSSALTCSMANCQYGCDVVKGQIRCQCPSPGLHLAPDGRTCDVDVDECATGRASCPFRQCVNT  
FGSYICKCHKGFDLMYIGGKYQCHDIDECSLGQYQCSSFARCYNVRGSYKCKCKEGYQGDGLT  
CVYIPKVMIEPSGPIHVPKNGTILKGDGTGNNNWIPDVGSTWWPPKTPYIIPPIITNRPTSKPT  
TRPTPKPTPIPTPPPPPLPTELRTPLPPTTPERPTTGLTTIAPAASTPPGGITVDNRVQTD  
QKPRGDFVSVLVHSCNFDHGLCGWIREKDNDLHWEPIRDPAGGQYLTVSAAKAPGGKAARLVL  
PLGRLMHSGDLCLSFHKVTGLHSGTLQVFVRKHGAHGAALWGRNGGHGWRQTQITLRGADIK  
SESQR

**Signal sequence:**

amino acids 1-17

**N-glycosylation site.**

amino acids 273-277

**Casein kinase II phosphorylation site.**

amino acids 166-170, 345-349

**Tyrosine kinase phosphorylation site.**

amino acids 199-206

**N-myristoylation site.**amino acids 109-115, 125-131, 147-153, 191-197, 221-227, 236-242,  
421-427, 433-439, 462-468, 476-482**Aspartic acid and asparagine hydroxylation site.**

amino acids 104-116, 186-198, 231-243

**Cell attachment sequence.**

amino acids 382-385

**EGF-like domain cysteine pattern signature.**

amino acids 75-87

53/615

**FIGURE 53**

CGGGCCGCCCCGGCCCCCATTCGGGCGGGGCTCGCTGCGGCGGGCGACTGAGCCAGGCTGGG  
CCGCGTCCCTGAGTCCCAGAGTCGGCGCGGCGCGGCAGGGGCAGCCTTCCACCACGGGGAGCC  
CAGCTGTCAGCCGCTCACAGGAAGATGCTGCGTGGGCGGGGCAGCCCTGGCATGGGTGTGCA  
TGTGGGTGCAGCCCTGGGAGCACTGTGGTTCTGCCTCACAGGAGCCCTGGAGGTCCAGGTCCC  
TGAAGACCCAGTGGTGGCACTGGTGGGCACCGATGCCACCCTGTGCTGCTCCTTCTCCCTGA  
GCCTGGCTTCAGCCTGGCACAGCTCAACCTCATCTGGCAGCTGACAGATACCAAACAGCTGGT  
GCACAGCTTTGCTGAGGGCCAGGACCAGGGCAGCGCCTATGCCAACCGCACGGCCCTCTTCCC  
GGACCTGCTGGCACAGGGCAACGCATCCCTGAGGCTGCAGCGCGTGCCTGTGGCGGACGAGGG  
CAGCTTCACCTGCTTCGTGAGCATCCGGGATTTTCGGCAGCGCTGCCGTGACCTGCAGGTGGC  
CGCTCCCTACTCGAAGCCAGCATGACCCTGGAGCCCAACAAGGACCTGCGGCCAGGGGACAC  
GGTGACCATCACGTGCTCCAGCTACCAGGGCTACCCTGAGGCTGAGGTGTTCTGGCAGGATGG  
GCAGGGTGTGCCCTGACTGGCAACGTGACCACGTGCGAGATGGCCAACGAGCAGGGCTTGTT  
TGATGTGCACAGCGTCCTGCGGGTGGTGTGGGTGCGAATGGCACCTACAGCTGCCTGGTGCG  
CAACCCCGTGCTGCAGCAGGATGCGCACRGCTCTGTCAACATCACAGGGCAGCCTATGACATT  
CCCCCAGAGGCCCTGTGGGTGACCGTGGGGCTGTCTGTCTGTCTCATTGCACTGCTGGTGGC  
CCTGGCTTTCGTGTGCTGGAGAAAGATCAAACAGAGCTGTGAGGAGGAGAATGCAGGAGCTGA  
GGACCAGGATGGGGAGGGAGAAGGCTCCAAGACAGCCCTGCAGCCTCTGAAACACTCTGACAG  
CAAAGAAGATGATGGACAAGAAATAGCC**TGAC**CCATGAGGACCAGGGAGCTGCTACCCCTCCCT  
ACAGCTCCTACCCTCTGGCTGCAATGGGGCTGCACTGTGAGCCCTGCCCCAACAGATGCATC  
CTGCTCTGACAGGTGGGCTCCTTCTCAAAGGATGCGATACACAGACCACTGTGCAGCCTTAT  
TTCTCCAATGGACATGATTCCCAAGTCATCCTGCTGCCTTTTTTCTTATAGACACAATGAACA  
GACCACCCACAACCTTAGTTCTCTAAGTCATCCTGCCTGCTGCCTTATTTACAGTACATACA  
TTTCTTAGGGACACAGTACACTGACCACATCACCACCCTCTTCTTCCAGTGCTGCGTGGACCA  
TCTGGCTGCCTTTTTTCTCCAAAAGATGCAATATTCAGACTGACTGACCCCTGCCTTATTT  
ACCAAAGACACGATGCATAGTCACCCCGGCCCTTGTTTCTCCAATGGCCGTGATACACTAGTGA  
TCATGTTTCAAGCCCTGCTTCCACCTGCATAGAATCTTTTCTTCTCAGACAGGGACAGTGCGGCC  
TCAACATCTCCTGGAGTCTAGAAGCTGTTTCTTCTCCCTCCTTCCCTGCCCCAAGTGAA  
GACAGGGCAGGGCCAGGAATGCTTTGGGGACACCGAGGGGACTGCCCCCACCACCATGG  
TGCTATTCTGGGGCTGGGGCAGTCTTTTCTGGCTTGCCCTCTGGCCAGCTCCTGGCCTCTGGT  
AGAGTGAGACTTCAGACGTTCTGATGCCTTCCGGATGTCATCTCTCCCTGCCCCAGGAATGGA  
AGATGTGAGGACTTCTAATTTAAATGTGGGACTCGGAGGGATTTTGTAACCTGGGGGTATATT  
TTGGGGAAAATAAATGTCTTTGTAAAAAAAAAAAAAAAAAAAAA



54/615

**FIGURE 54**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA41386
><subunit 1 of 1, 316 aa, 1 stop, 1 unknown
><MW: -1, pI: 4.62, NX(S/T): 4
MLRRRGSPGMGVHVGAAALGALWFCLTGALEVQVPEDPVVALVGTDATLCCSFSPPEPGFSLAQL
NLIWQLTDTKQLVHSFAEGDQGSAYANRTALFPDLLAQGNASLRLQVRVADGSGFTCFVSI
RDFGSAAVSLQVAAPYSKPSMTLEPNKDLRPGDVTITCSSYQGYPEAEVFWQDGQGVPLTGN
VTTSQMANEQGLFDVHSVLRVVLGANGTYSCLVRNPVLQQDAHXSVTITGQPMTFPPEALWVT
VGLSVCLIALLLVALAFVCWRKIKQSCEEENAGAEDQDGEGEKSKTALQPLKHSDSKEDDGQEIA
```

**Important features:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 251-270

**N-glycosylation site.**

amino acids 91-94, 104-107, 189-192 and 215-218

**Homologous region to Immunoglobulins and MHC**

amino acids 217-234

55/615

**FIGURE 55**

GAGTCTTGACCGCCGCGGGCTCTTGGTACCTCAGCGCGAGCGCCAGGCGTCCGGCCGCGCGTG  
GCT**ATG**TTCGTGTCCGATTTCCGCAAAGAGTTCTACGAGGTGGTCCAGAGCCAGAGGGTCCTT  
CTCTTCGTGGCCTCGGACGTGGATGCTCTGTGTGCGTGCAAGATCCTTCAGGCCTTGTTCCAG  
TGTGACCACGTGCAATATACGCTGGTTCCAGTTTCTGGGTGGCAAGAAGTTGAAACTGCATTT  
CTTGAGCATAAAGAACAGTTTCATTATTTTATTCTCATAAACTGTGGAGCTAATGTAGACCTA  
TTGGATATTCTTCAACCTGATGAAGACACTATATTCTTTGTGTGTGACTCCCATAGGCCAGTC  
AATGTCGTCAATGTATACAACGATACCCAGATCAAATTACTCATTAACAAGATGATGACCTT  
GAAGTTCCCGCCTATGAAGACATCTTCAGGGATGAAGAGGAGGATGAAGAGCATTTCAGGAAAT  
GACAGTGATGGGTCAGAGCCTTCTGAGAAGCGCACACGGTTAGAAGAGGAGATAGTGGAGCAA  
ACCATGCGGAGGAGGCAGCGGCGAGAGTGGGAGGCCCGGAGAAGAGACATCCTCTTTGACTAC  
GAGCAGTATGAATATCATGGGACATCGTCAGCCATGGTGATGTTTGAGCTGGCTTGGATGCTG  
TCCAAGGACCTGAATGACATGCTGTGGTGGGCCATCGTTGGACTAACAGACCAGTGGGTGCAA  
GACAAGATCACTCAAATGAAATACGTGACTGATGTTGGTGTCTGCAGCGCCACGTTTCCCGC  
CACAACCACCGGAACGAGGATGAGGAGAACACACTCTCCGTGGACTGCACACGGATCTCCTTT  
GAGTATGACCTCCGCCTGGTGCTCTACCAGCACTGGTCCCTCCATGACAGCCTGTGCAACACC  
AGCTATACCGCAGCCAGGTTCAAGCTGTGGTCTGTGCATGGACAGAAGCGGCTCCAGGAGTTC  
CTTGACAGACATGGGTCTTCCCCTGAAGCAGGTGAAGCAGAAGTTCCAGGCCATGGACATCTCC  
TTGAAGGAGAATTTGCGGGAAATGATTGAAGAGTCTGCAAATAAATTTGGGATGAAGGACATG  
CGCGTGACAGACTTTCAGCATTTCATTTTGGGTTCAAGCACAAAGTTTCTGGCCAGCGACGTGGTC  
TTTGCCACCATGTCTTTGATGGAGAGCCCCGAGAAGGATGGCTCAGGGACAGATCACTTCATC  
CAGGCTCTGGACAGCCTCTCCAGGAGTAACCTGGACAAGCTGTACCATGGCCTGGAACCTCGCC  
AAGAAGCAGCTGCGAGCCACCCAGCAGACCATTGCCAGCTGCCTTTGCACCAACCTCGTCATC  
TCCCAGGGGCCTTTTCTGTACTGCTCTCTCATGGAGGGCACTCCAGATGTCATGCTGTTCTCT  
AGGCCGGCATCCCTAAGCCTGCTCAGCAAACACCTGCTCAAGTCCTTTGTGTGTTTCGACAAAG  
AACC GGCGCTGCAAAGTCTGCCCCCTGGTGATGGCTGCCCCCTGAGCATGGAGCATGGCACA  
GTGACCGTGGTGGGCATCCCCCAGAGACCGACAGCTCGGACAGGAAGAAGTTTGGGAGG  
GCGTTTGAGAAGGCAGCGGAAAGCACCAGCTCCCGGATGCTGCACAACCATTGACCTCTCA  
GTAATTGAGCTGAAAGCTGAGGATCGGAGCAAGTTTCTGGACGCACTTATTTCCCTCCTGTCC  
**TAG**GAAATTTGATTCTTCCAGAATGACCTTCTTATTTATGTAAGTGGCTTTTCAATTTAGATTGTA  
AGTTATGGACATGATTTGAGATGTAGAAGCCATTTTTTTATTAAATAAAATGCTTATTTTAGGAAA

56/615

**FIGURE 56**

MFVSDFRKEFYEVVQSQRVLLFVASDVDALCACKILQALFQCDHVQYTLVPVSGWQELET AFL  
EHKEQFHYFILINCGANVDLLDILQPD EDTIFFVCDSHRPVNVVNVYNDTQIKLLIKQDD DLE  
VPAYEDI FRDEEEDEEHSGNDS DGSEPSEK RTRLEEEIVEQTMRRRQRREWEARRRDILFDYE  
QY EYHGTSSAMVMFELAWMLSKDLNDMLWWAIVGLTDQWVQDKITQMKYVTDVGV LQRHVSRH  
NHRNEDEENTLSVDCTRISFEYDLRLVLYQHWSLHDSL CNTSYTAARFKLWSVHGQKRLQEFL  
ADMGLPLKQVKQKFQAMDISLKENLREMIEESANKFGMKDMRVQTF SIHFGFKHKFLASDVVF  
ATMSLMESPEKDGSGTDHFIQALDSLRSNLDKLYHGLELAKKQLRATQQT IASCLCTNLVIS  
QGPFLYCSLMEGTPDVMLFSRPASLSLLSKHLLKS FVCSTKNRRCKLLPLVMAAPLSMEHGT V  
TVVGIPPETDSSDRKNFFGRAFEKAAESTSSRMLHNHFDLSVIELKAEDRSKFLDALISLLS

57/615

**FIGURE 57**

CGCCGCCGTTGGGGCTGGAAGTTCCCGCCAGGTCCGTGCCGGGCGAGAGAGATGCTGCCCGGC  
CCGCCTCGGCTTTGAGGCGAGAGAAGTGTCCAGACCCATTTGCGCTTGCTGACGGCGTTCGAG  
CCCTGGCCAGAC**ATGT**CCACAGGGTTCTCCTTCGGGTCCGGGACTCTGGGCTCCACCACCGTG  
GCCGCCGGCGGGACAGCACAGGCGGCGTTTTCTCCTTCGGAACGGGAACGTCTAGCAACCCT  
TCTGTGGGGCTCAATTTTGGAAATCTTGGAAGTACTTCAACTCCAGCAACTACATCTGCTCCT  
TCAAGTGGTTTTTGGAAACGGGCTCTTTGGATCTAAACCTGCCACTGGGTTCACTCTAGGAGGA  
ACAAATACAGGTGCCTTGACACCAAGAGGCCTCAAGTGGTCACCAAATATGGAACCCTGCAA  
GGAAACAGATGCATGTGGGGAAGACACCCATCCAAGTCTTTTTAGGAGTCCCCTTCTCCAGA  
CCTCCTCTAGGTATCCTCAGGTTTGCACCTCCAGAACCCCGGAGCCCTGGAAAGGAATCAGA  
GATGCTACCACCTACCCGCTGGATGGAGTCTCGCTCTGTGCCAGGCTGGAGTGCAGTGGCA  
CGATCTCGGCTCACTGCAACCTCCGCTCCCGGGTTCAAGCGAGTCTCCTGCCTCAGCCTCTG  
AGTGTCTGGGGCTACAGGTGCCTGCAGGAGTCCCTGGGGCCAGCTGGCCTCGATGTACGTCAGC  
ACGCGGGAACGGTACAAGTGGCTGCGCTTCAGCGAGGACTGTCTGTACCTGAACGTGTACGCG  
CCGGCGCGCGCGCCCGGGGATCCCCAGCTGCCAGTGTGGTCTGGTTCCCGGGAGGCGCCTTC  
ATCGTGGGCGCTGCTTCTTCGTACGAGGGCTCTGACTTGGCCGCGCCGAGAAAGTGGTGCTG  
GTGTTTTCTGCAGCACAGGCTCGGCATCTTCGGCTTCCTGAGCACGGACGACAGCCACGCGCGC  
GGAACTGGGGGCTGCTGGACCAGATGGCGGCTCTGCGCTGGGTGCAGGAGAACATCGCAGCC  
TTCGGGGGAGACCCAGGAAATGTGACCCTGTTTCGGCCAGTCGGCGGGGGCCATGAGCATCTCA  
GGACTGATGATGTACCCCTAGCCTCGGGTCTCTTCCATCGGGCCATTTCCAGAGTGGCACC  
GCGTTATTCAACTTTTCATCACTAGTAACCCACTGAAAGTGGCCAAGAAGGTTGCCACCTG  
GCTGGATGCAACCACAACAGCACACAGATCCTGGTAAACTGCCTGAGGGCACTATCAGGGACC  
AAGGTGATGCGTGTGTCCAACAAGATGAGATTCTCCAACCTGAACCTCCAGAGAGACCCGGAA  
GAGATTATCTGGTCCATGAGCCCTGTGGTGGATGGTGTGGTGTATCCAGATGACCCTTTGGTG  
CTCCTGACCCAGGGGAAGGTTTTCATCTGTGCCCTACCTTCTAGGTGTCAACAACCTGGAATTC  
AATTGGCTCTTGCTTATAATATACCAAGGAGCAGGTACCACTTGTGGTGGAGGAGTACCTG  
GACAATGTCAATGAGCATGACTGGAAGATGCTACGAAACCGTATGATGGACATAGTTCAAGAT  
GCCACTTTCGTGTATGCCACACTGCAGACTGCTCACTACCACCGAGAAACCCCAATGATGGGA  
ATCTGCCCTGCTGGCCACGCTACAACAAGGATGAAAAGTACCTGCAGCTGGATTTTACCACAA  
GAGTGGGCAT**TGA**AGCTCAAGGAGAAGAAGATGGCTTTTTGGATGAGTCTGTACCAGTCTCAA  
GACCTGAGAAGCAGAGGCAATTCTAAGGGTGGCTATGCAGGAAGGAGCCAAAGAGGGGTTTGC  
CCCCACCATCCAGGCCCTGGGGAGACTAGCCATGGACATACCTGGGGACAAGAGTTCTACCCA  
CCCCAGTTTAGAACTGCAGGAGCTCCCTGCTGCCTCCAGGCCAAAGCTAGAGCTTTTGCCTGT  
TGTGTGGGACCTGCACTGCCCTTTCAGCCTGACATCCATGATGCCCTCTACTTCACTGTT  
GACATCCAGTTAGGCCAGGCCCTGTCAACACCACACTGTGCTCAGCTCTCCAGCCTCAGGACA  
ACCTCTTTTTTTCCCTTCTTCAAATCCTCCCACCCTCAATGTCTCCTTGTGACTCCTTCTTA  
TGGGAGGTCGACCCAGACTGCCACTGCCCCGTGCTACTGCACCCAGCTTGGCATTTACCATCCA  
TCCTGCTCAACCTTGTTCCTGTCTGTTACATTGGCCTGGAGGCCTAGGGCAGGTTGTGACAT  
GGAGCAAACCTTTTGGTAGTTTGGGATCTTCTCTCCCACCCACACTTATCTCCCCAGGGCCAC  
TCCAAAGTCTATACACAGGGGTGGTCTCTTCAATAAAGAAGTGTGATTAGAAAAA

58/615

**FIGURE 58**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44179
<subunit 1 of 1, 545 aa, 1 stop
<MW: 58934, pI: 9.45, NX(S/T): 4
MSTGFSFGSGTLGSTTVAAGGTSTGGVFSFGTGTSSNPSVGLNFGNLGSTSTPATTSA PSSGF
GTGLFSGSKPATGFTLGSTNTGALHTKRPQVVTKYGT LQGKQMHVGKTPIQVFLGVPF SRPPLG
ILRFAPPEPPEPWKGIRDATTYPGWSLALSPGWSA VARSRLTATSASRVQASLLPQPLSVWG
YRCLQESWGQLASMYVSTRERYKWLRFSEDCLYLNVYAPARAPGDPQLPVMVWFPGGAFIVGA
ASSYEGSDLAAREKVVLVFLQHRLGIFGFLSTDDSHARGNWGLLDQMAALRWVQENIAAFGGD
PGNVTLFGQSAGAMSISGLMMSPLASGLFHRAISQSGTALFRLFITSNPLKVAKKVAHLAGCN
HNSTQILVNCLRALSGTKVMRVS NKMRFQLNFQRPDEEIIWSMSPVVDGVVIPDDPLVLLTQ
GKVSSVPYLLGVNNLEFNWLLPYNITKEQVPLVVEEYLDNVNEHDWKMLRNRMMDIVQDATFV
YATLQTAHYHRETPMMGICPAGHATTRMKSTCSWILPQEWA
```

**Important features:****Signal peptide:**

amino acids 1-29

**Carboxylesterases type-B serine active site.**

amino acids 312-327

**Carboxylesterases type-B signature 2.**

amino acids 218-228

**N-glycosylation sites.**

amino acids 318-321, 380-383 and 465-468



60/615

**FIGURE 60**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA44192

&lt;subunit 1 of 1, 694 aa, 1 stop

&lt;MW: 77400, pI: 9.54, NX(S/T): 6

MLLLGLCLGLSLCVGSQEEAQSWGHSSEQDGLRVPRQVRLQLKTKPLMTEFSVKSTIISR  
YAFTTVSCRMLNRASEDQDIEFQMQIPAAAFITNFTMLIGDKVYQGEITEREKKSGDRVKEKR  
NKTTEENGEKGTETFRASAVIPSKDKAAFFLSYEELLQRRLGKYEHSISVRPQQLSGRLSVDV  
NILESAGIASLEVLPLHNSRQRGSGRGEDDSGPPPSTVINQNETFANIIIFKPTVVQQARIAQN  
GILGDFIIRYDVNREQSIGDIQVLNGYFVHYFAPKDLPLPKNVFVLDSSASMVGTKLRQTK  
DALFTILHDLRPQDRFSIIGFSNRIKVWKDHLISVTPDSIRDGKVYIHHMSPTGGTDINGALQ  
RAIRLLNKYVAHSGIGDRSVSLIVFLTDGKPTVGETHTLKILNNTREAARGQVCIFTIGIGND  
VDFRLLLEKLSLENCGLTRRVHEEEDAGSQLIGFYDEIRTPLLSDIRIDYPPSSVVQATKTLFP  
NYFNGSEIIAGKLVDRKLDHLHVEVTASNSKKFIILKTDVPVRPQKAGKDVGTGSPRGGDGE  
GDTNHIERLWSYLTTKELLSSWLQSDDEPEKERLRQRAQALAVSYRFLTPFTSMKLRGPVPRM  
DGLEEAHGMSAAMGPEPVVQSVRGAGTQPGPLLKKPNSVKKKQNKTKKRHRDGVFPLHHLGIR

**Signal sequence.**

amino acids 1-14

**N-glycosylation sites.**

amino acids 97-101, 127-131, 231-235, 421-425, 508-512, 674-678

**Glycosaminoglycan attachment sites.**

amino acids 213-217, 391-395

**N-myristoylation sites.**

amino acids 6-12, 10-16, 212-218, 370-376, 632-638, 638-644

61/615

**FIGURE 61**

CAGGAACCCCTCTCTTTGGGTCTGGATTGGGACCCCTTTCCAGTACCATTTTTTCTAGTGAACC  
ACGAAGGGACGATACCAGAAAAACACCCTCAACCCAAAGGAAATAGACTACAGCCCCAATTGGC  
TGACTTTGGCTATAGAAAAAAGAAAGGAACGAAAAGAGACAGTTTTTTTTTGGAAAGCTAAGTC  
TTCCCTTTATCGAGTCAAGAAACCCCCCTTCTTGAGCTATTTACAGCTTTTAACAATTGAGT  
AAAGTACGCTCCGGTCAACC**ATGGT**GACAGCCGCCCTGGGTCCCGTCTGGGCAGCGCTCCTGCT  
CTTTCTCCTGATGTGTGAGATCCGTATGGTGGAGCTCACCTTTGACAGAGCTGTGGCCAGCGG  
CTGCCAACGGTGTGTGACTCTGAGGACCCCTGGATCCTGCCCATGTATCCTCAGCCTCTTC  
CTCCGGCCGCCCCACGCCCTGCCTGAGATCAGACCCTACATTAATATCACCATCCTGAAGGG  
TGACAAAGGGGACCCAGGCCCAATGGGCCTGCCAGGGTACATGGGCAGGGAGGGTCCCCAAGG  
GGAGCCTGGCCCTCAGGGCAGCAAGGGTGACAAGGGGGAGATGGGCAGCCCCGGCGCCCCGTG  
CCAGAAGCGCTTCTTCGCCTTCTCAGTGGGCCGCAAGACGGCCCTGCACAGCGCGAGGACTT  
CCAGACGCTGCTCTTCGAAAGGGTCTTTGTGAACCTTGATGGGTGCTTTGACATGGCGACCGG  
CCAGTTTGCTGCTCCCCCTGCGTGGCATCTACTTCTTCAGCCTCAATGTGCACAGCTGGAATTA  
CAAGGAGACGTACGTGCACATTATGCATAACCAGAAAGAGGCTGTATCCTGTACGCGCAGCC  
CAGCGAGCGCAGCATCATGCAGAGCCAGAGTGTGATGCTGGACCTGGCCTACGGGGACCGCT  
CTGGGTGCGGCTCTTCAAGCGCCAGCGCGAGAAGCCCATCTACAGCAACGACTTCGACACCTA  
CATCACCTTCAGCGGCCACCTCATCAAGGCCGAGGACGACT**GA**GGGCCTCTGGGCCACCCTCC  
CGGCTGGAGAGCTCAGGTGCTGGTCCCGTCCCTGCAGGGCTCAGTTTGCACTGCTGTGAAGC  
AGGAAGGCCAGGGAGGTCCCCGGGGACCTGGCATTTCTGGGGAGACCCTGCTTCTATCTTGGCT  
GCCATCATCCCTCCCAGCCTATTTCTGCTCCTCTCTTCTCTCTTGGACCTATTTTAAGAAGCT  
TGCTAACCTAAATATTCTAGAACTTTCCAGCCTCGTAGCCCAGCACTTCTCAAACCTTGGAAA  
TGATGGGAATCACCCGGGGTTCGTGTTAAATGCAGATTCTGACTCAGCAGGTCTGAGTGGGT  
CCAGGATTCTGTGTTTCTCATATGTTCTGGGTGATGCTGATGGGGTCAGTCTATGAACCACA  
CTGGAGCAACCAGGTTCTAGGACTTTCTCAATATTCTAGTACTTTCTGAACATTCTGGAATCC  
TCCCCACATTCTAGAATTCTCCCAACATTTTTTTTTTCTTGAGACAGAGTCTTGCTCTGTTGCC  
CAGGTAGAGTGCAAGTGGTGCATCTCAGTTCACTGCAACCTCTGCCTCCCGGGTTCAAGCGA  
TTCTTCTGCCTCAGCCTCCCTAGTGGCTGGGATTACAGGCGCCTGCTACCATGCCTGGCTAAT  
TTTTGTATTTTATAGTAGAGATGGGGTTTACCATATTGGCCAGGCTGGTCTTGAACCTCCTGAC  
TTCAGGTGACCCACCCGCTCGGCCTCTCAAAATGCTGGGATTACAGGTGTGAGCCACCGTGC  
CTGGCCAATTCCAACATTCTTAAATTTCTCTCATCCCTCCAGGGCTCCCCGTGCTATGTTCTCT  
TTACCCCTTCCCCCTCTTCTCTTGCTCAGGCCTGCACCACTGCAGCCACCGTTTCAATTTATTCA  
TTCATTAAACACTGAGCACTCACTCTGTGCTGGGTCCCGGGAAGGGTGAGGGGGTCAGACACA  
GGCCCTGCCCCTGCCCTCAGTGACTGGCCAGTCCAGCCCAGGCGGGGAGAGATGTGTACATAG  
GTTTTAAAGCAGACCCAGAGCTCATGGGGGCTGTGTTCTGGGTGTTTCAAGGTGCTGCTGGTCC  
TCCATTACCCACTGCTCCCCAAGGCTGGTGGGACGGGGTCCCGGTGGCAGGGGAGGTATCTC  
CTTCCCGTTCCCTCATCCACCTGCCAGTGCTCATCGTTACAGCAAACCCAGGGGGCTTGGC  
CAGGTCAAGGGTTCTGTGAGGAGAGGACCCAGGAGTGTGGGGGCATTTGGGGGGTGAAGTGGC  
CCCCGAAGAATGGAACCCACACCCATAGCTCTCCCCACAGCTGATACGGCATCCTGCGAGAAG  
ACCTGCCCTCCTCACTGGGATCCCCCTTCCCTGCCTCCTCCCAGGGCTCTGCCAGGGCCTTGCTC  
AGTCCCTTCCACCAAAGTCATCTGAACCTCCGTTTCCCCAGGGCCTCCAGCTGCCCTCAGACA  
CTGATGTCTGTCCCCAGGTGCTCTCTGCCCTCATGCCCCCTCTACCCGGCCAGTGCCCCGAC  
TCTCCAGGCTTTATCAAGGTGCTAAGGCCCGGGTGGGCAGCTCCTCGTCTCAGAGCCCTCCTC  
CGGCCTGGTGCTGCCTTTACAAACACCTGCAGGAGAAGGGCCACGGAAGCCCCAGGCTTTAGA  
GCCCTCAGCAGGTCTGGGGAGCTAGAGCAAAGGAGGGACCTCAGGCCTTCCGTTTCTTCTTCC  
AGGGTGGGGTGGCCTGGTGTTCCTTAGCCTTCCAAACCCAGGTGGCCTGCCCTTCTCCCCAG  
AGGGAGGCGGCCTCCGCCATTGGTGCTCATGCAGACTCTGGGGCTGAGGTGCCCGGGGGGT  
GATCTCTGGTGCTCACAGCCGAGGGAGCCGTGGCTCCATGGCCAGATGACGGAACAGGGTCT  
GACCAAGTGCCAGGAAGACCTGTGCTATAAACCACCCTGCCTGATCCTGCCCTGCCTGACCC  
CGCCACGCCCTGCCGTCCAGCATGATTAAAGAATGCTGTCTCCTCTTGAAAAA



62/615

**FIGURE 62**

MVTAALGPVWAALLLFLLMCEIRMVELTFDRAVASGCQRCCDSEDPLDPAHVSSASSSSGRPHA  
LPEIRPYINITILKGDKGDPGPMGLPGYMGREGPQGEPPGQSGDKGEMGSPGAPCQKRFFA  
FSVGRKTALHSGEDFQTLLFERVFVNLDGCFDMATGQFAAPLRGIYFFSLNVHSWNYKETVH  
IMHNQKEAVILYAQPSESRIMQSQSVMLDLAYGDRVWVRLFKRQRENAIYSNDFDTYITFSGH  
LIKAEDD

**Important features:**

**Signal peptide:**

amino acids 1-20

**N-glycosylation site.**

amino acids 72-75

**Clq domain proteins.**

amino acids 144-178, 78-111 and 84-117

63/615

**FIGURE 63**

ATGGAAGCCAGTAACACTGTGGCCTACTATCTCTTCCGTGGTGCCATCTACATTTTTTGGGAC  
TCGGGAATTATGAGGTAGAGGTGGAGGCGGAGCCGGATGTCAGAGGTCCTGAAATAGTCACCA  
**TGGGGGAAAATGATCCGCCTGCTGTTGAAGCCCCCTTCTCATTCCGATCGCTTTTTTGGCCTTG**  
ATGATTTGAAAATAAGTCCTGTTGCACCAGATGCAGATGCTGTTGCTGCACAGATCCTGTCAC  
TGCTGCCATTGAAGTTTTTTCCAATCATCGTCATTGGGATCATTGCATTGATATTAGCACTGG  
CCATTGGTCTGGGCATCCACTTCGACTGCTCAGGGAAGTACAGATGTCGCTCATCCTTTAAGT  
GTATCGAGCTGATAGCTCGATGTGACGGAGTCTCGGATTGCAAAGACGGGGAGGACGAGTACC  
GCTGTGTCCGGGTGGGTGGTCAGAATGCCGTGCTCCAGGTGTTACAGCTGCTTCGTGGAAGA  
CCATGTGCTCCGATGACTGGAAGGGTCACTACGCAAATGTTGCCTGTGCCAACTGGGTTTTCC  
CAAGCTATGTGAGTTCAGATAACCTCAGAGTGAGCTCGCTGGAGGGGGCAGTTCCGGGAGGAGT  
TTGTGTCCATCGATCACCTCTTGCCAGATGACAAGGTGACTGCATTACACCACTCAGTATATG  
TGAGGGAGGGATGTGCCTCTGGCCACGTGGTTACCTTGCACTGCACAGCCTGTGGTCATAGAA  
GGGGCTACAGCTCACGCATCGTGGGTGGAAACATGTCCTTGCTCTCGCAGTGGCCCTGGCAGG  
CCAGCCTTCAGTTCCAGGGCTACCACCTGTGCGGGGGCTCTGTCAACACGCCCTGTGGATCA  
TCACTGCTGCACACTGTGTTTATGACTTGTACCTCCCCAAGTCATGGACCATCCAGGTGGGTC  
TAGTTTTCCCTGTTGGACAATCCAGCCCCATCCCACTGGTGGAGAAGATTGTCTACCACAGCA  
AGTACAAGCCAAAGAGGGCTGGGCAATGACATCGCCCTTATGAAGCTGGCCGGGGCCACTCACGT  
TCAATGAAATGATCCAGCCTGTGTGCCTGCCCAACTCTGAAGAGAACTTCCCCGATGGAAAAG  
TGTGCTGGACGTCAGGATGGGGGGCCACAGAGGATGGAGGTGACGCCTCCCCGTCTCTGAACC  
ACGCGGGCCGTCCCTTTGATTTCCAACAAGATCTGCAACCACAGGGACGTGTACGGTGGCATCA  
TCTCCCCCTCCATGCTCTGCGCGGGCTACCTGACGGGTGGCGTGGACAGCTGCCAGGGGGACA  
GCGGGGGGGCCCCTGGTGTGTCAAGAGAGGAGGCTGTGGAAGTTAGTGGGAGCGACCAGCTTTG  
GCATCGGCTGCGCAGAGGTGAACAAGCCTGGGGTGTACACCCGTGTCACCTCCTTCTGGACT  
GGATCCACGAGCAGATGGAGAGAGACCTAAAAACCT**TGA**AGAGGAAGGGGACAAGTAGCCACCT  
GAGTTCCCTGAGGTGATGAAGACAGCCCGATCCTCCCCCTGGACTCCCGTGTAGGAACCTGCACA  
CGAGCAGACACCCTTGGAGCTCTGAGTTCGGGCACCAGTAGCAGGCCCCGAAAGAGGCACCCTT  
CCATCTGATTCCAGCACAACCTTCAAGCTGCTTTTTTGTTTTTTGTTTTTTTGAGGTGGAGTCT  
CGCTCTGTTGCCCAGGCTGGAGTGCAGTGGCGAAATCCCTGCTCACTGCAGCCTCCGCTTCCC  
TGTTTCAAGCGATTCTCTTGCCCTCAGCTTCCCCAGTAGCTGGGACCACAGGTGCCCGCCACCA  
CACCCAATAATTTTTGTATTTTGTAGTAGAGACAGGGTTTTACCATGTTGGCCAGGCTGCTCT  
CAAACCCCTGACCTCAAATGATGTGCCTGCTTCAGCCTCCCACAGTGTGGGATTACAGGCAT  
GGGCCACCACGCCTAGCCTCACGCTCCTTTCTGATCTTCACTAAGAACAAAAGAAGCAGCAAC  
TTGCAAGGGCGGCCTTTCCCACTGGTCCATCTGGTTTTCTCTCCAGGGTCTTGCAAAATTCCT  
GACGAGATAAGCAGTTATGTGACCTCACGTGCAAAGCCACCAACAGCCACTCAGAAAAGACGC  
ACCAGCCCAGAAGTGCAGAACTGCAGTCACTGCACGTTTTTCATCTCTAGGGACCAGAACCAAA  
CCCACCCTTTCTACTTCCAAGACTTATTTTCACATGTGGGGAGGTTAATCTAGGAATGACTCG  
TTTAAGGCCTATTTTCATGATTTCTTTGTAGCATTTGGTGCTTGACGTATTATGTCTTTGA  
TTCCAAATAATATGTTTCCTTCCCTCATTGTCTGGCGTGTCTGCGTGGACTGGTGACGTGAAT  
CAAAATCATCCACTGAAA

64/615

**FIGURE 64**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45234
><subunit 1 of 1, 453 aa, 1 stop
><MW: 49334, pI: 6.32, NX(S/T): 1
MGENDPPAVEAPFSFRSLFGLDDLKISPVAPDADAVAAQILSLLPLKFFPIIVIGIIALILAL
AIGLGIHFDCSGKYRCRSSFKCIELIARCDGVSDCKDGEDEYRCVRVGGQNAVQVFTAASWK
TMCSDDWKGYANVACAQLGFPSYVSSDNLRVSSLEGQFREEFVSDHLLPDDKVLTALHHSVY
VREGCASGHVVTLQCTACGHRRGYSSRIVGGNMSLLSQWPWQASLQFQGYHLCGGSVITPLWI
ITAAHCVDLYLPKSWTIQVGLVSLLDNPAPSHLVEKIVYHSHYKPKRLGNDIALMKLAGPLT
FNEMIQPVCLPNSEENFPDGKVCWTSGWGATEDGGDASPVLNHAAPPLISNKICNHRDVYGGI
ISPSMLCAGYLTGGVDSCQGDSGGPLVCQERRLWKLVGATSEFGIGCAEVNKPGRVYTRVTSFLD
WIHEQMERDLKT
```

**Signal Peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 240-284

65/615

**FIGURE 65**

CGGGCCAGCCTGGGGCGGCCGGCCAGGAACCAACCCGTTAAGGTGTCTTCTCTTTAGGGATGGT  
GAGGTTGGAAAAAGACTCCTGTAACCCTCCTCCAGGATGAACCACTGCCAGAAGACATGGAG  
AACGCTCTCAACGGGAGCCAGAGCTCCCATGCTTCTCTGCGCAATATCCATTCCATCAACCCC  
ACACAACCTCATGGCCAGGATTGAGTCCTATGAAGGAAGGGAAAAGAAAGGCATATCTGATGTC  
AGGAGGACTTTCTGTTTGTGTTGTCACCTTTGACCTCTTATTCGTAACATTACTGTGGATAATA  
GAGTTAAATGTGAATGGAGGCATTGAGAACACATTAGAGAAGGAGGTGATGCAGTATGACTAC  
TATTCTTCATATTTTGATATATTTCTTCTGGCAGTTTTTCGATTTAAAGTGTTAATACTTGCA  
TATGCTGTGTGCAGACTGCGCCATTGGTGGGCAATAGCGTTGACAACGGCAGTGACCAGTGCC  
TTTTTACTAGCAAAAGTGATCCTTTTGAAGCTTTTCTCTCAAGGGGCTTTTGGCTATGTGCTG  
CCCATCATTTTCATTCATCCTTGCTGGATTGAGACGTGGTTCCTGGATTTCAAAGTGTTACCT  
CAAGAAGCAGAAGAAGAAAACAGACTCCTGATAGTTCAGGATGCTTCAGAGAGGGGCAGCACTT  
ATACCTGGTGGTCTTTCTGATGGTCAGTTTTATTCCCTCCTGAATCCGAAGCAGGATCTGAA  
GAAGCTGAAGAAAAACAGGACAGTGAGAAACCACTTTTAGAACTATGAGTACTACTTTTGTTA  
AATGTGAAAAACCCTCACAGAAAGTCATCGAGGCAAAAAGAGGCAGGCAGTGAGTCTCCCTG  
TCGACAGTAAAGTTGAAATGGTGACGTCCACTGCTGGCTTTATTGAACAGCTAATAAAGATTT  
ATTTATTGTAATACCTCACAAACGTTGTACCATATCCATGCACATTTAGTTGCCTGCCTGTGG  
CTGGTAAGGTAATGTCATGATTCATCCTCTCTTCAGTGAGACTGAGCCTGATGTGTTAACAAA  
TAGGTGAAGAAAGTCTTGTGCTGTATTCCCTAATCAAAAGACTTAATATATTGAAGTAACACTT  
TTTTAGTAAGCAAGATACCTTTTTTATTTCAATTCACAGAATGGAATTTTTTTGTTTCATGTCT  
CAGATTTATTTTGTATTTCTTTTTTAACACTCTACATTTCCCTTGTTTTTTAACTCATGCACA  
TGTGCTCTTTGTACAGTTTTTAAAAAGTGTAATAAAATCTGACATGTCAATGTGGCTAGTTTTTA  
TTTTTCTTGTTTTGCATTATGTGTATGGCCTGAAGTGTTGGACTTGCAAAAGGGGAAGAAAGG  
AATTGCGAATACATGTAAATGTCACCAGACATTTGTATTATTTTTATCATGAAATCATGTTT  
TTCTCTGATTGTTCTGAAATGTTCTAAATACTCTTATTTTGAATGCACAAAATGACTTAAACC  
ATTCATATCATGTTTTCTTTGCGTTCAGCCAATTTCAATTAAAATGAACTAAATTAAAAA

66/615

**FIGURE 66**

MNHLPEDMENALTGSQSSHASLRNIHSINPTQLMARIESYEGREKKGISDVRRTFCLFVTFDL  
LFVTLLWIIELNVNGGIENTLEKEVMQYDYYSSYFDIFLLAVFRFKVLILAYAVCRLRHWWAI  
ALTTAVTSAFLLAKVILSKLFSQGAFGYVLPPIISFILAWIETWFLDFKVLQPQEAEEENRLLIV  
QDASERAALIPGGLSDGQFYSPPESEAGSEEAEKQDSEKPLLEL

**Important features of the protein:****Signal peptide:**

amino acids 1-20

**Transmembrane domains:**

amino acids 54-72, 100-118, 130-144, 146-166

**N-myristoylation sites.**

amino acids 14-20, 78-84, 79-85, 202-208, 217-223

67/615

**FIGURE 67**

AATAAAGCTTCCTTAATGTTGTATATGTCTTTGAAGTACATCCGTGCATTTTTTTTTTAGCATC  
CAACCATTCCTCCCTTGTAGTTCTCGCCCCCTCAAATCACCTCTCCCGTAGCCACCCGACT  
AACATCTCAGTCTCTGAAA**ATG**CACAGAGATGCCTGGCTACCTCGCCCTGCCTTCAGCCTCAC  
GGGGCTCAGTCTCTTTTTCTTTTGGTGCCACCAGGACGGAGCATGGAGGTCACAGTACCTGC  
CACCTCAACGTCTCAATGGCTCTGACGCCCGCCTGCCCTGCACCTTCAACTCCTGCTACAC  
AGTGAACCACAAACAGTTCTCCCTGAAGTGGACTTACCAGGAGTGCAACAAGTCTCTGAGGA  
GATGTTCTCCAGTTCCGCATGAAGATCATTAACTGAAGCTGGAGCGGTTTTCAAGACCGCGT  
GGAGTTCTCAGGGAACCCAGCAAGTACGATGTGTGGTGATGCTGAGAAAGTGCAGCCGGA  
GGATGAGGGGATTTACAAGTCTACATCATGAACCCCCCTGACCGCCACCGTGGCCATGGCAA  
GATCCATCTGCAGGTCTCATGGAAGAGCCCCCTGAGCGGGACTCCACGGTGGCCGTGATTGT  
GGGTGCCCTCCGTGGGGGCTTCTGGCTGTGGTCTCTTGGTGCTGATGGTGGTCAAGTGTGT  
GAGGAGAAAAAAGAGCAGAAGCTGAGCACAGATGACCTGAAGACCGAGGAGGAGGGCAAGAC  
GGACGGTGAAGGCAACCCGGATGATGGCGCCAAG**TAG**TGGGTGGCCGGCCCTGCAGCCTCCCG  
TGTCCCGTCTCTCCCTCTCCGCCCTGTACAGTGACCTGCCTGCTCGCTCTTGGTGTGCTT  
CCCGTGACCTAGGACCCAGGGCCACCTGGGGCCTCTGAACCCCCGACTTCGTATCTCCCA  
CCCTGCACCAAGAGTGACCCACTCTCTCCATCCGAGAACTGCCATGCTCTGGGACGTGTG  
GGCCCTGGGGAGAGGAGAGAAAGGGCTCCACCTGCCAGTCCCTGGGGGAGGCAGGAGGCAC  
ATGTGAGGGTCCCCAGAGAGAAGGGAGTGGGTGGGCAGGGGTAGAGGAGGGGCCGTGTACC  
TGCCAGTGTCTGCCTGGCAGTGGCTTCAGAGAGGACCTGGTGGGGAGGGAGGGCTTTCCTGT  
GCTGACAGCGCTCCCTCAGGAGGGCCTTGGCCTGGCACGGCTGTGCTCCTCCCTGCTCCAG  
CCCAGAGCAGCCATCAGGCTGGAGGTGACGATGAGTTCTTGAAGTGGAGGGGCATGTTAAA  
GGGATGACTGTGCATTCCAGGGCACTGACGGAAGCCAGGGCTGCAGGCAAAGCTGGACATGT  
GCCCTGGCCCAGGAGGCCATGTTGGGCCCTCGTTTCCATTGCTAGTGGCCTCCTTGGGGCTCC  
TGTTGGCTCCTAATCCCTTAGGACTGTGGATGAGGCCAGACTGGAAGAGCAGCTCCAGGTAGG  
GGGCCATGTTTCCCAGCGGGGACCCACCAACAGAGGCCAGTTTCAAAGTCAGCTGAGGGGCTG  
AGGGGTGGGGCTCCATGGTGAATGCAGGTGCTGCAGGCTCTGCCTTCTCCATGGGGTAACCA  
CCCTCGCCTGGGCAGGGGCAGCCAAGGCTGGGAAATGAGGAGGCCATGCACAGGGTGGGGCAG  
CTTCTTTGGGGCTTCAGTGAGAACTCTCCAGTTGCCCTTGGTGGGGTTTCCACCTGGCTTT  
TGGCTACAGAGAGGGAAGGGAAAGCCTGAGGCCGGCATAAGGGGAGGCCTTGGAACTGAGCT  
GCCAATGCCAGCCCTGTCCCATCTGCGGCCACGCTACTCGCTCCTCTCCCAACAACTCCCTTC  
GTGGGGACAAAAGTGACAATTGTAGGCCAGGCACAGTGGCTCACGCCTGTAATCCCAGCACTT  
TGGGAGGCCAAGGCGGGTGGATTACCTCCATCTGTTTAGTAGAAATGGGCAAAACCCCATCTC  
TACTAAAAATACAAGAATTAGCTGGGCGTGGTGGCGTGTGCCTGTAATCCCAGCTATTTGGGA  
GGCTGAGGCAGGAGAATCGCTTGAGCCCGGAAGCAGAGGTTGCAGTGAAGTGAATAGTAGTAT  
AGTGCCACTGCAATTCAGCCTGGGTGACATAGAGAGACTCCATCTCAAAAAAA

68/615

**FIGURE 68**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45415

&lt;subunit 1 of 1, 215 aa, 1 stop

&lt;MW: 24326, pI: 6.32, NX(S/T): 4

MHRDAWLPRPAFSLTGLSLFFSLVPPGRSMEVTVPATLNVLNGSDARLPCTFNSCYTVNHKQF  
SLNWTYQECNNCSEEMFLQFRMKIINLKLERFQDRVEFSGNPSKYDVSVMLRNVQPEDEGIYN  
CYIMNPPDRHRGHGKIHLQVLMEEPPERDSTVAVIVGASVGGFLAVVILVLMVVKCVRRKKEQ  
KLSTDDLKTEEEGKTDGEGNPDDGAK

**Important features:****Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 161-179

**Immunoglobulin-like fold:**

amino acids 83-127

**N-glycosylation sites.**

amino acids 42-45, 66-69 and 74-77

69/615

**FIGURE 69**

GGCGCCTGGTTCTGCGGCTACTGGCTGTACGGAGCAGGAGCAAGAGGTCGCCGCCAGCCTCCGCCGCCGAGCCTC  
GTTTCGTGTCCCCGCCCTCGCTCCTGCAGCTACTGCTCAGAAACGCTGGGGCGCCACCCTGGCAGACTAACGAA  
GCAGCTCCCTTCCCACCCCAACTGCAGGTCTAATTTTGGACGCTTTGCCTGCCATTCTTCCAGGTTGAGGGAGC  
CGCAGAGGCGGAGGCTCGCGTATTCTGTCAGTCAGCACCCACGTCGCCCGCGGACGCTCGGTGCTCAGGCCCTTC  
GCGAGCGGGGCTCTCCGTCTGCGGTCCCTTGTGAAGGCTCTGGGCGGCTGCAGAGGCCGCGCGTCCGGTTTGGCT  
CACCTCTCCCAGGAACTTCACACTGGAGAGCCAAAAGGAGTGGAAGAGCCTGTCTTGAGATTTTCTGGGGAA  
ATCCTGAGGTCATTCAATTATGAAGTGTACCGCGCGGGAGTGGCTCAGAGTAACCACAGTGCTGTTTCATGGCTAGA  
GCAATTCCAGCCATGGTGGTCCCAATGCCACTTTATTGGAGAACTTTTGGAAAAATACATGGATGAGGATGGT  
GAGTGGTGGATAGCCAAACAACGAGGGAAAAGGGCCATCACAGACAATGACATGCAGAGTATTTTGGACCTTCAT  
AATAAATTACGAAGTCAGGTGTATCCAACAGCCTCTAATATGGAGTATATGACATGGGATGTAGAGCTGGAAAGA  
TCTGCAGAATCCTGGGCTGAAAGTTGCTTGTGGGAACATGGACCTGCAAGCTTGCTTCCATCAATTGGACAGAAT  
TTGGGAGCACACTGGGGAAGATATAGGCCCCGACGTTTTCATGTACAATCGTGGTATGATGAAGTGAAGACTTT  
AGCTACCCATATGAACATGAATGCAACCCATATTGTCCATTGAGGTGTTCTGGCCCTGTATGTACACATTATACA  
CAGGTCTGTGGGCAACTAGTAACAGAATCGGTTGTGCCATTAATTTGTGTACATAACATCTGGGGGCGAG  
ATATGGCCCCAAGCTGTCTACCTGGTGTGCAATTACTCCCCAAGGGAACTGGTGGGGCCATGCCCCCTTACAAA  
CATGGGCGGCCCTGTTCTGCTTGCCCCCTAGTTTTTGGAGGGGGCTGTAGAGAAAATCTGTGCTACAAAGAAGGG  
TCAGACAGGTATTATCCCCCTCGAGAAGAGGAAACAAATGAAATAGAACGACAGCAGTCACAAGTCCATGACACC  
CATGTCCGGACAAGATCAGATGATAGTAGCAGAAATGAAGTCATAAGCGCACAGCAAATGTCCCAAATTGTTTCT  
TGTGAAGTAAGATTAAAGAGATCAGTGCAAGGAACAACCTGCAATAGGTACGAATGTCCTGCTGGCTGTTTGGAT  
AGTAAAGCTAAAGTTATTGGCAGTGTACATTATGAAATGCAATCCAGCATCTGTAGAGCTGCAATTCATTATGGT  
ATAATAGACAATGATGGTGGCTGGGTAGATATCACTAGACAAGGAAGAAAGCATTATTTTCATCAAGTCCAATAGA  
AATGGTATTCAAACAATTGGCAAATATCAGTCTGCTAATTCCTTCACAGTCTCTAAAGTAACAGTTCAGGCTGTG  
ACTTGTGAAACAACCTGTGGAACAGCTCTGTCCATTTTATAAGCCTGCTTCACATTGCCCCAAGAGTATACTGTCCCT  
CGTAACTGTATGCAAGCAAATCCACATTATGCTCGTGTAAATTGGAACCTCGAGTTTATTCTGATCTGTCCAGTATC  
TGCAGAGCAGCAGTACATGCTGGAGTGGTTCGAAATCACGGTGGTTATGTTGATGTAATGCCTGTGGACAAAAGA  
AAGACCTACATTGCTTCTTTTTCAGAAATGGAATCTTCTCAGAAAGTTTACAGAATCCTCCAGGAGGAAAGGCATTC  
AGAGTGTGTTGCTGTTGTGTGAAGTGAATACTTGAAGAGGACCATAAAGACTATTCCAAATGCAATATTTCTGA  
ATTTTGTATAAAACTGTAACTTACTGTACAGAGTACATCACTATTTTCAGCCCCAAAAGGTGCCAAATGCATA  
TAAATCTTGATAAACAAAGTCTATAAAATAAAACATGGGACATTAGCTTTGGGAAAAGTAATGAAAATATAATGG  
TTTTAGAAATCCTGTGTTAAATATTGCTATATTTTCTTAGCAGTTATTTCTACAGTTAATTACATAGTCATGATT  
GTTCTACGTTTCATATATTATATGGTGCCTTGTATATGCCACTAATAAAATGAATCTAAACATTGAATGTGAATG  
GCCCTCAGAAAATCATCTAGTGCATTTAAAAATAATCGACTCTAAACTGAAAGAAACCTTATCACATTTTCCCC  
AGTTCAATGCTATGCCATTACCAACTCCAAATAATCTCAAATAATTTTCCACTTAATAACTGTAAAGTTTTTTTC  
TGTTAATTTAGGCATATAGAATATTAAATCTGATATTGCACTTCTTATTTTATATAAAATAATCCTTTAATATC  
CAAATGAATCTGTAAATGTTTGATTCTTGGGAATGGCCTTAAAAATAAATGTAATAAAGTCAGAGTGGTGGT  
ATGAAAACATTCCTAGTGATCATGTAGTAAATGTAGGGTTAAGCATGGACAGCCAGAGCTTTCTATGTACTGTTA  
AAATTGAGGTCACATATTTTCTTTTGTATCCTGGCAAATACTCCTGCAGGCCAGGAAGTATAATAGCAAAAAGTT  
GAACAAAGATGAACATAATGTATTACATTACCATTGCCACTGATTTTTTTTAAATGGTAAATGACCTTGATATAA  
ATATTGCCATATCATGGTACCTATAATGGTGATATATTTGTTTCTATGAAAAATGTATTGTGCTTTGATACTAAA  
AATCTGTAAATGTAGTTTTGGTAATTTTTTTCTGCTGGTGGATTACATATTAAATTTTTTCTGCTGGTGGGA  
TAAACATTAAATTAATCATGTTTCAAAAAAAAAAAAAA



70/615

**FIGURE 70**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45417

&lt;subunit 1 of 1, 500 aa, 1 stop

&lt;MW: 56888, pI: 8.53, NX(S/T): 2

MKCTAREWLRVTTVLFMARAI PAMVVPNATLLEKLLEKYMDEDGEWWIAKQRGKRAITDNDMQ  
SILDLHNKLR SQVYPTASNMEYMTWDVELERSAESWAESCLWEHGPASLLPSIGQNLGAHWGR  
YRPPTFHVQSWYDEVKDFSYPYEHECNPYCPFRCSGPVCTHYTQV VWATSNRIGCAINLCHNM  
NIWGQIWPKAVYLV CNYSPKGNWWGHAPYKHGRPCSACPPSFGGGCRENLCYKEGSDRYYPFR  
EETNEIERQQSQVHDTHVRTRSDSSRNEVIS AQQMSQIVSCEVRLRDQCKGTTCNRYECPA  
GCLDSKAKVIGSVHYEMQSSICRAAIHYGIIDNDGGWVDITRQGRKHYFIKSNRNGIQTIGKY  
QSANSFTVSKVTVQAVTCETTVEQLCPFHKPASHCPRVYCPRNCMQANPHYARVIGTRVYSDL  
SSICRAAVHAGVVRNHGGYVDVMPVDKRKTYIASFQNGIFSES LQNPPGGKAFRVFAVV

**Important features:****Signal peptide:**

amino acids 1-20

**Extracellular proteins SCP/Tpx-1/Ag5/PR-1/Sc7 protein**

amino acids 165-186, 196-218, 134-146, 96-108 and 58-77

**N-glycosylation site**

amino acids 28-31

71/615

**FIGURE 71**

CAGCCCCGCGCGCCGGCCGAGTCGCTGAGCCGCGGCTGCCGGACGGGACGGGACCGGCTAGGC  
TGGGCGCGCCCCCGGGCCCCGCGTGGGC**ATG**GGCGCACTGGCCCCGGGCGCTGCTGCTGCCT  
CTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCC  
CTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCC  
GAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCC  
GCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAG  
ATGCTGATCGGGACCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTT  
GCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACA  
TACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTT  
GGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT  
ATTTTTGAATCAGAGAATTTCTTTTTGCGTGGGATTAAATGGAATGGAATACTTGGCCTAGCT  
TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAA  
GCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCT  
GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATC  
TGGTATAACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGA  
GGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGC  
ACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAAGCTGTGGCCCGCGCA  
TCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAAT  
TCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGACGAGAACTCCAGC  
AGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTACAGCCCATGATGGGGGCCGGCCTG  
AATTATGAATGTTACCGATTTCGGCATTTCCTCCATCCACAAATGCGCTGGTGATCGGTGCCACG  
GTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC  
CCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGAT  
GTAGCCAGCAACTGTGTCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCTAT  
GCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTT  
CGGTGTCAGCGTCGCCCCGTGACCCTGAGGTCTCAATGATGAGTCCTCTCTGGTCAGACAT  
CGCTGGAA**TGA**ATAGCCAGGCCTGACCTCAAGCAACCATGAACTCAGCTATTAAGAAAATCA  
CATTTCCAGGGCAGCAGCCGGGATCGATGGTGGCGCTTCTCCTGTGCCACCCGTCTTCAAT  
CTCTGTTCTGCTCCCAGATGCCTTCTAGATTCACTGTCTTTTGATTCTTGATTTTCAAGCTTT  
CAAATCCTCCCTACTTCCAAGAAAAATAATTAAAAAAAACCTTCATTCTAA

72/615

**FIGURE 72**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA45493
><subunit 1 of 1, 518 aa, 1 stop
><MW: 56180, pI: 5.08, NX(S/T): 2
MGALARALLLPLLAQWLLRAAPELAPAPFTLPLRVAAATNRVVAPTGGPGTPAERHADGLALA
LEPALASPAGAAANFLAMVDNLQGDSEGRGYYLEMLIGTPPQKLQILVDTGSSNFAVAGTPHSYI
DTYFDTERSSTYRSKGFDTVYKYTQGSWTGFVGEDLV TIPKGENTSFLVNIATIFESENFFLP
GIKWNGILGLAYATLAKPSSSLETFFDSLVTQANIPNVFSMQMCGAGLPVAGSGTNGGSLVLG
GIEPSLYKGDWYTPIKEEWYYQIEILKLEIGGQSLNLD CREYNADKAIVD SGTTLLRLPQKV
FDAVVEAVARASLIPEFSDGFWTGSQACWTNSETPWSYFPKISIIYLRDENSSRSFRITILPQ
LYIQPMMGAGLNYECYRFGISPSTNALVIGATVMEGFYVIFDRAQKRVGFAASPCAEIAGAAV
SEISGPFSTEDVASNCVPAQSLSEPILWIVSYALMSVCGAILLV LIVLLLLPFRQRRPRDPE
VVNDESSLVRHRWK
```

**Important features:****Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 466-494

**N-glycosylation sites.**

amino acids 170-173 and 366-369

**Leucine zipper pattern.**

amino acids 10-31 and 197-118

**Eukaryotic and viral aspartyl proteases**

amino acids 109-118, 252-261 and 298-310

73/615

**FIGURE 73**

GCCGCGGCGAGAGCGCGCCAGCCCCGCCGCGATGCCGCGCGCCAGGACGCCTCCTCCCGCTGCTGGCCCCGGC  
CGGCGGCCCTGACTGCGCTGCTGCTGCTGCTGCTGGGCCATGGCGGCGCGGGCGCTGGGGCGCCCGGGCCAGG  
AGGCGGCGGCGGCGGCGGCGGACGGGCCCCCGCGGCGAGACGGCGAGGACGGACAGGACCCGCACAGCAAGCACC  
TGTACACGGCCGACATGTTACGCGACGGGATCCAGAGCGCCGCGCACTTCGTTCATGTTCTTCGCGCCCTGGTGTG  
GACACTGCCAGCGGCTGCAGCCGACTTGAATGACCTGGGAGACAAATACAACAGCATGGAAGATGCCAAAGTCT  
ATGTGGCTAAAGTGGACTGCACGGCCCACTCCGACGTGTGCTCCGCCAGGGGGTGGGAGGATACCCACCTTAA  
AGCTTTTCAAGCCAGGCCAAGAAGCTGTGAAGTACCAGGGTCCTCGGGACTTCCAGACACTGGAAAACTGGATGC  
TGCAGACACTGAACGAGGAGCCAGTGACACCAGAGCCGGAAGTGAACCGCCAGTGCCCCGAGCTCAAGCAAG  
GGCTGTATGAGCTCTCAGCAAGCAACTTTGAGCTGCACGTTGCACAAGGCGACCACTTTATCAAGTTCTTCGCTC  
CGTGGTGTGGTCACTGCAAAGCCCTGGCTCCAACCTGGGAGCAGCTGGCTCTGGGCCTTGAACATTCGAAACTG  
TCAAGATTGGCAAGGTTGATTGTACACAGCACTATGAACTCTGCTCCGGAACACAGGTTCTGGCTATCCCACTC  
TTCTCTGGTTCCGAGATGGGAAAAAGGTGGATCAGTACAAGGGAAAGCGGGATTTGGAGTCACTGAGGGAGTACG  
TGGAGTCGCAGCTGCAGCGCACAGAGACTGGAGCGACGGAGACCGTCACGCCCTCAGAGGCCCGGTGCTGGCAG  
CTGAGCCCCGAGGCTGACAAGGGCACTGTGTTGGCACTCACTGAAAAATAACTTCGATGACACCATTGCAGAAGGAA  
TAACCTTCATCAAGTTTTATGCTCCATGGTGTGGTCACTGTAAGACTCTGGCTCCTACTTGGGAGGAACTCTCTA  
AAAAGGAATTCCCTGGTCTGGCGGGGTCAAGATCGCCGAAGTAGACTGCACTGCTGAACGGAATATCTGCAGCA  
AGTATTCGGTACGAGGCTACCCACGTTATTGCTTTCCGAGGAGGGAAGAAAGTCAGTGAGCACAGTGGAGGCA  
GAGACCTTGACTCGTTACACCGCTTTGTCTGAGCCAAGCGAAAGACGAACCTTAGGAACACAGTTGGAGGTCAC  
CTCTCCTGCCAGCTCCCGCACCCCTGCGTTTAGGAGTTCACTCCACAGAGGCCACTGGGTTCCAGTGGTGGCT  
GTTCAGAAAAGCAGAACATACTAAGCGTGAGGTATCTTCTTTGTGTGTGTGTTTTCCAAGCCAACACACTCTACAG  
ATTCTTTATTAAGTTAAGTTTCTCTAAGTAAATGTGTAACCTCATGGTCACTGTGTAACATTTTTCAGTGGCGATA  
TATCCCTTTGACCTTCTCTTGATGAAATTTACATGGTTTTCTTTGAGACTAAAATAGCGTTGAGGGAAATGAAA  
TTGCTGGACTATTTGTGGCTCCTGAGTTGAGTGATTTTGGTGAAAGAAAGCACATCCAAAGCATAGTTTACCTGC  
CCACGAGTTCGGAAGGTGGCCTTGTGGCAGTATTGACGTTCTCTGATCTTAAGGTACAGTTGACTCAATAC  
TGTGTTGGTCCGTAGCATGGAGCAGATTGAAATGCAAAAACCCACACCTCTGGAAGATACCTTCACGGCCGCTGC  
TGGAGCTTCGTTGCTGTGAATACTTCTCTCAGTGTGAGAGTTAGCCGTGATGAAAGCAGCGTTACTTCTGACC  
GTGCCTGAGTAAGAGAATGCTGATGCCATAACTTTATGTGTGATACCTTGTCAAATCAGTTACTGTTTCAGGGGAT  
CCTTCTGTTTCTCACGGGTGAAACATGTCTTTAGTTTCTCATGTTAACACGAAGCCAGAGCCACATGAACTGT  
TGGATGTCTTCTTAGAAAGGGTAGGCATGGAAAATTCCACGAGGCTCATTCTCAGTATCTCATTAACTCATTGA  
AAGATTCCAGTTGTATTTGTACCTGGGGTGACAAGACCAGACAGGCTTTCCAGGCCTGGGTATCCAGGGAGGC  
TCTGCAGCCCTGCTGAAGGGCCCTAACTAGAGTTCTAGAGTTTCTGATTCTGTTTCTCAGTAGTCTTTTAGAGG  
CTTGCTATACTTGGTCTGCTTCAAGGAGGTCGACCTTCTAATGTATGAAGAATGGGATGCATTTGATCTCAAGAC  
CAAAGACAGATGTGAGTGGGCTGCTCTGGCCCTGGTGTGCACGGCTGTGGCAGCTGTTGATGCCAGTGTCTCTA  
ACTCATGCTGTCTTGTGATTAAACACCTCTATCTCCCTTGGGAATAAGCACATACAGGCTTAAGCTCTAAGATA  
GATAGGTGTTTGTCTTTTACCATCGAGCTACTTCCATAATAACCACTTTGCATCCAACACTCTTACCCACCT  
CCCATACGCAAGGGGATGTGGATACTTGGCCCAAAGTAACTGGTGGTAGGAATCTTAGAAACAAGACCACTTATA  
CTGTCTGTCTGAGGCAGAAGATAACAGCAGCATCTCGACCAGCCTCTGCCTTAAAGGAAATCTTTATTAATCAG  
TATGGTTACAGATAATTCTTTTTTTAAAAAAACCCCAACCTCCTAGAGAAGCACAACTGTCAAGAGTCTTGTA  
CACAACCTTCAGCTTTGCATCAGAGTCTTGTATTCCAAGAAAATCAAAGTGGTACAATTTGTTTGTACACTAT  
GATACTTTCTAAATAAACTCTTTTTTTTTAA

74/615

**FIGURE 74**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA46776
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47629, pI: 5.90, NX(S/T): 0
MPARPGRLLPLLARPAALTALLLLLLGHGGGRWGARAQEAADGPPAADGEDGQDPHSK
HLYTADMFTHGIQSAAHFVMFFAPWCGHCQRLQPTWNDLGDKYNSMEDAKVYVAKVDCTAHS
VCSAQGVRGYPTLKLKFKPGQEAVKYQGPRDFQTLNWMQLTLNEEPVTPEPEVEPPSAPELKQ
GLYELASNFELHVAQGDHFIKFFAPWCGHCKALAPTWEQLALGLEHSETVKIGKVDCTQHVE
LCSGNQVRGYPTLLWFRDGKKVDQYKGRDLESLREYVESQLQRTETGATETVTPSEAPVLAA
EPEADKGTVLALTENNFDFTIAEGITFIKFYAPWCGHCKTLAPTWEELSKKEFPGLAGVKIAE
VDCTAERNICSKYSVRGYPTLLLFRGGKKVSEHSGGRDLDSLHFRVLSQAKDEL
```

**Signal sequence:**  
amino acids 1-32

75/615

**FIGURE 75A**

CGGACGCGTGGGCGGACGCGTGGGCAAAAGAACTCGGAGTGCCAAAGCTAAATAAGTTAGCTGAGAAAACGCACG  
CAGTTTGCAGCGCTTGGCCGGGTGCGCCAACTACGCAAAGACCAAGCGGGCTCCGCGCGGACCGGCCGCGGGC  
TAGGGACCCGGCTTTGGCCCTTACAGGCTCCCTAGCAGCGGGGAAAAGGAATTGCTGCCCGGAGTTTCTGCGGAGGT  
GGAGGGAGATCAGGAAACGGCTTCTTCTCACTTCGCGCGCTGGTGAGTGTGCGGGAGATTGGCAAACGCCTAGG  
AAAGGACTGGGGAAAATAGCCCTGGGAAAGTGGAGAAGGTGATCAGGAGGCCGGTCCACTACGGCAGTTTATCTG  
TCTGATCAGAGCCAGACGCGACGCGTCCACTTCGCAAGTTCTTTCCAGGTGTGGGGACCGCAGGACAGACGGCCGA  
TCCCGCCGCCCTCCGTACCAGCACTCCCAGGAGAGTCAGCCTCGCTCCCCAACGTGAGGGCGCTCTGGCCACGA  
AAAGTTCTGTCCACTGTGATTCTCAATTCCTTGCTTGGTTTTTTTCTCCAGAGAAGTTTGGGTGGAGATATTA  
ACTTTTTTCTTTTTTTTTTCTTGGTGGAAGCTGCTCTAGGGAGGGGGAGGAGGAGGAGAAAGTGAAATGTGC  
TGGAGAAGAGCGAGCCCTCCTTGTCTTCCGGAGTCCCATTCCATTAAGCCATCACTTCTGGAAGATTAAAGTTGT  
CGGACATGGTGACAGCTGAGAGGAGAGGAGGATTTCTTGCCAGGTGGAGAGTCTTACCGTCTGTTGGGTGCATG  
TGTGCGCCCGCAGCGGCGCGGGCGCGTGGTTCTCCGCGTGAGTCTCACCTGGGACCTGAGTGAATGGCTCCCA  
GGGGCTGTFGCGGGGCATCCGCTCCGCTTCTCCACAGGCCTGTGTCTGTCTGGAAGATGCTAGCAATGGGGG  
CGCTGGCAGGATTCGGATCCTCTGCCTCCTCACTTATGGTTACCTGTCTGGGGCCAGGCCTTAGAAGAGGAGG  
AAGAAGGGGCCCTTACTAGCTCAAGCTGGAGAGAACTAGAGCCAGCACAACCTCCACCTCCAGCCCCATCTCA  
TTTTCATCTAGCGGATGATCAGGGATTTAGAGATGTGGGTACCACGGATCTGAGATTAACACCTACTCTTG  
ACAAGCTCGCTGCCGAAGGAGTTAAACTGGAGAAGTACTATGTCCAGCCTATTTGCACACCATCCAGGAGTCAGT  
TTATTACTGGAAAGTATCAGATACACACCGGACTTCAACATCTATCATAAGACCTACCAACCCCACTGTTTAC  
CTCTGGACAATGCCACCCTACCTCAGAACTGAAGGAGGTTGGATATTCAACGCATATGGTCGGAAAATGGCACT  
TGGGTTTTTAACAGAAAAGATGCATGCCCACCAGAAGAGGATTTGATACCTTTTTTGGTTCCCTTTTGGGAAGTG  
GGGATTACTATACACACTACAAATGTGACAGTCTGGGATGTGTGGCTATGACTTGTATGAAAACGACAATGCTG  
CCTGGGACTATGACAATGGCATATACTCCACACAGATGTACACTCAGAGAGTACAGCAAATCTTAGCTTCCCATA  
ACCCACAAAGCCTATATTTTTATATACTGCCTATCAAGCTGTTCACTCACCAGTCAAGCTCCTGGCAGGTATT  
TCGAACACTACCGATCCATTATCAACATAAACAGGAGAAGATATGCTGCCATGCTTTCCTGATTGAAGCAA  
TCAACACAGTGAACCTTGGCTCTAAAGACTTATGGTTTCTATAACAACAGCATTATCATTTACTCTTCAGATAATG  
GTGGCCAGCCTACGGCAGGAGGGAGTAAGTGGCTCTCAGAGGTAGCAAAGGAACATATTGGGAAGGAGGGATCC  
GGGCTGTAGGCTTTGTGCATAGCCCACTTCTGAAAAACAAGGGAACAGTGTGTAAGGAACCTGTGCACATCACTG  
ACTGGTACCCCACTCTCATTTCACTGGCTGAAGGACAGATTGATGAGGACATTCAACTAGATGGCTATGATATCT  
GGGAGACCATAAGTGAGGGTCTTCCGCTCACCCGAGTAGATATTTGCATAACATTGACCCCTATACACCAAGGC  
AAAAATGGCTCCTGGGCAGCAGGCTATGGGATCTGGAACACTGCAATCCAGTCAGGCATCAGAGTCAGCACTG  
GAAATTGCTTACAGGAAATCCTGGCTACAGCGACTGGGTCCCCCTCAGTCTTTAGCAACCTGGGACCGAACC  
GTGGCACAATGAACGGATCACCTTGTCAACTGGCAAAAGTGTATGGCTTTCAACATCACAGCCGACCCATATGA  
GAGGGTGGACCTATCTAACAGGTATCCAGGAATCGTGAAGAAGCTCCTACGAGGCTCTCACAGTTCAACAAAC  
TGCAGTGCCGTCAGGTATCCCCCAAAGACCCAGAAAGTAACCTAGGCTCAATGGAGGGGTCTGGGGACCATG  
GTATAAAGAGGAAACCAAGAAAAAGCAAGCAAAATCAGGCTGAGAAAAAGCAAAAGAAAAAGCAAAAAAA  
GAAGAAGAAACAGCAGAAAGCAGTCTCAGGTAAACAGCAAAATTTGGCTCGATAATATCGCTGGCCTAAGCGTCA  
GGCTTGTTTTTCATGCTGTGCCACTCCAGAGACTTCTGCCACCTGGCGCCACACTGAAACCTGCTCTGCTCAGT  
CCAAGGTGCTACTCTTGAAGCCACACTTAGAGAGAGTGGAGATGTTTATTTCTCTCGCTCCTTTAGAAAACGTG  
GTGAGTCTGAGTCCACTGCTGTGCTTCAGTCAACTGACCAACACTGCTTTGAATTATAGGAGGAGAACATA  
ACCTACCATCCGCAAGCATGCTAATTTGATGGAAGTTACAGGGTAGCATGATTAACCTACCTTTGATAAATTAC  
AGTCAAAGATTGTGTACCTCAAAGGCCTTGAAGAATATATTTCTTGGTGAATTTTTGTATGTCTGTATATGA  
CACTTGGGTTTTTTAATTAATCTATTTTATATATAAATATATGTTTCTTTTCTGTGAAAAGCTGTTTTCT  
CACATGTGAACAGCTTGCACCTCATTTTACCATCGCTGAGGGAATGGCAAATAAGAATGTTTGAGCACACTGCCC  
ACAATGAATGTAACATTTTTCTAAACACTTTACTAGAAGAACATTTAGTATAAAAAACCTAATTTATTTTACA  
GAAAAATATTTTGTGTTTTTATAAAAAGTTATGCAATGACTTTTATTTTTATTTCTGCATACCATTAGAAGA  
ATTTTATTTTCAATTTCTCAAATTATCAAGCACTGTAATACTATAAATTAATGTAATACTGTGTGAATTCAGACTA  
TAAAAACATCATTACAGAAACCTTATAATCGTCATTGTTCAATCAAGATTTTGAATGTAATAAGATGAATATAT  
ATTACTTGGAAATTCAATGTTTGTGCAGAGTTGAGACAACCTTTATGTTTCTATCATAAACTATTTATGTATCTT  
AATTATTAATGATTACTTTATGGCACTAGAAATTTACTGTGGCTTTTCTGATCTAATCTTAGCTTAAAT  
GTATCATGTTCTCTAAAAAATAAAAACTTTTACTAATAGGCAATGGAAGGAATGGTTTGTCTAACACCCACAGTAA  
TATAATATGATTTTACAGATAGATGCTTCCCCTGGCTATGACATGGAGAAAGATTTTCCCATAATAAATACTAA  
TATTTATATTAGGTTGGTGCAAACTAGTTGCGGTTTTTCCCATTAAAAGTAATAACCTTACTCTATACAAAGT  
GGACACTGTGGGGAGATACAGAGAAATGGAAGATACGGATCCTGCCTGGAGTAGGTAACCTTGCTTGGAAACCC  
ACATGCAACGTGATGAGGAGAAATTAAGGAGTATTATCAGTAATGAAGTTTATCATGGGTGATCAATGAGCATA  
GATTGGTGTGGATCCTGTAGACCCTGGTGTCTTTGAAGTGCCCTCTCCTAATGCAGAGGCCTTGAAGCTTAC

FIGURE 75B

[illegible]

77/615

**FIGURE 76**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48296
><subunit 1 of 1, 515 aa, 1 stop
><MW: 56885, pI: 6.49, NX(S/T): 5
MAPRGCAGHPPPPSPQACVCPGKMLAMGALAGFWILCLLTGYLSWGQALEEEEEEGALLAQAGEKLEPSTTSTSQ
PHLIFILADDQGFRDVGYPYHGSEIKTPTLDKLAAGVVKLENYYVQPICTPSRSQFITGKYQIHTGLQHSIIRPTQP
NCLPLDNATLPQKLKEVGYSTHMGKWHLGFNKECMPTRRGFDTEFGSLLGSGDYTHYKCDSPGMCQGYDLYEN
DNAAWDYDNGIYSTQMYTORVQQILASHNPTKPIFLYTAYQAVHSPLQAPGRYFEHYRSIININRRRYAAMLSC
DEAINNVTLALKTYGFYNNSSIIYSSDNGGQPTAGGSNWPLRGSKGTYWEGGIRAVGFVHSPLLKNKGTVCKELV
HITDWYPTLISLAEGQIDEDIQLDGYDIWETISEGLRSPRVLDILHNIDPYTPRQKMAPGQQAMGSGTLQSSQPSE
CSTGNCLQEILATATGSPLSLSATWDRTGGTMNGSPCQLAKVYGFSTSQPTHMRGWTYLTGIQES
```

**Important Features:****Signal Peptide:**

amino acids 1-37

**Sulfatases signature 1.**

amino acids 120-132

**Sulfatases signature 2.**

amino acids 168-177

**Tyrosine kinase phosphorylation site.**

amino acids 163-169

**N-glycosylation sites.**

amino acids 157-160, 306-309 and 318-321



78/615

**FIGURE 77**

AAAAAAGCTCACTAAAGTTTCTATTAGAGCGAATACGGTAGATTTCCATCCCCTTTTGAAGAACAGTACTGTGGA  
GCTATTTAAGAGATAAAAACGAAATATCCTTTCTGGGAGTTCAAGATTGTGCAGTAATTGGTTAGGACTCTGAGC  
GCCGCTGTTACCAATCGGGGAGAGAAAAGCGGAGATCCTGCTCGCCTTGCACGCGCCTGAAGCACAAAGCAGAT  
AGCTAGGAATGAACCATCCCTGGGAGTATGTGGAACAACGGAGGAGCTCTGACTTCCCAACTGTCCCCATTCTAT  
GGGCGAAGGAACTGCTCCTGACTTCAGTGGTTAAGGGCAGAATTGAAAATAATTCTGGAGGAAGATAAGAAATGAT  
TCCTGCGCGACTGCACCGGGACTACAAAGGGCTTGTCTGCTGGGAATCCTCCTGGGGACTCTGTGGGAGACCGG  
ATGCACCCAGATACGCTATTCAGTTCGGGAAGAGCTGGAGAAAGGCTCTAGGCTGGGCGACATCTCCAGGGACCT  
GGGGCTGGAGCCCCGGGAGCTCGCGGAGCGCGGAGTCCGCATCATCCCCAGAGGTAGGACGCAGCTTTTCGCCCT  
GAATCCGCGCAGCGGCAGCTTGGTCACGGCGGGCAGGATAGACCGGGAGGAGCTCTGTATGGGGGCCATCAAGTG  
TCAATTAATCTAGACATTCTGATGGAGGATAAAGTGAAAATATATGGAGTAGAAGTAGAAGTAAGGGACATTAA  
CGACAATGCGCCTTACTTTTCGTGAAAGTGAATTAGAAAATAAAAATTAGTGAAAATGCAGCCACTGAGATGCGGTT  
CCCTCTACCCACCGCCTGGGATCCGGATATCGGGAAGAACTCTCTGCAGAGCTACGAGCTCAGCCCGAACACTCA  
CTTCTCCCTCATCGTGCAAAATGGAGCCGACGGTAGTAAGTACCCGAATTGGTGCTGAAACGCGCCCTGGACCG  
CGAAGAAAAGGCTGCTCACCACCTGGTCCTTACGGCCTCCGACGGGGCGGACCCGGTGCGCACAGGCACCGCGCG  
CATCCGCGTGATGGTTCTGGATGCGAACGACAACGCACCCAGCGTTTGTCTACGCCGAGTACCGCGCGAGCGTTCC  
GGAGAATCTGGCCTTGGGCACGCAGCTGCTTGTAGTCAACGCTACCGACCTGACGAAGGAGTCAATGCGGAAGT  
GAGGTATTCTTCCGGTATGTGGACGACAAGGCGGCCCAAGTTTTCAAAGTATGTAATTGAGGACAAATATC  
AACAATAGGGGAGTTGGACCACGAGGAGTCAAGATTCTACCAGATGGAAGTGCAAGCAATGGATAATGCAGGATA  
TTCTGCGCGAGCCAAAGTCTGATCACTGTTCTGGACGTGAACGACAATGCCCCAGAAGTGGTCTCACCTCTCT  
CGCCAGCTCGGTTCCCGAAAACCTCTCCAGAGGGACATTAATTGCCCTTTTAAATGTAAATGACCAAGATTCTGA  
GGAAAACGGACAGGTGATCTGTTTCATCCAAGGAAATCTGCCCTTTAAATTAGAAAAATCTTACGGAAATTACTA  
TAGTTTAGTCACAGACATAGTCTTGGATAGGGAACAGGTTCTTAGCTACAACATCACAGTGACCGCCACTGACCG  
GGGAACCCCGCTTATCCACGGAACCTCATATCTCGCTGAACGTGGCAGACACCAACGACAACCCCGCGTCTT  
CCCTCAGGCCTCCTATTCGGCTTATATCCAGAGAACAATCCAGAGGAGTTTCCCTCGTCTCTGTGACCGCCCA  
CGACCCCGACTGTGAAGAGAACGCCCAGATCACTTATTCCTGGCTGAGAACACCATCCAAGGGGCAAGCCTATC  
GTCTACGTGTCCATCAACTCCGACACTGGGGTACTGTATGCGCTGAGCTCCTTCGACTACGAGCAGTTCGAGGA  
CTTGCAAGTGAAAGTGATGGCGCGGGACAACGGGCACCCGCCCTCAGCAGCAACGTGTCTGTTGAGCCTGTTCTGT  
GCTGGACCAGAACGACAATGCGCCGAGATCCTGTACCCCGCCCTCCCAACGGACGGTTCCACTGGCGTGGAGCT  
GGTCCCCGCTCCGAGAGCCCGGTACCTGGTGACCAAGGTGGTGGCGGTGGACAGAGACTCCGCGCCAGAACGC  
CTGGCTGTCTACCGTCTGCTCAAGGCCAGCGAGCCGGGACTCTTCTCGGTGGGTCTGCACACGGGCGAGGTGCG  
CACGGCGCGAGCCCTGCTGGACAGAGACGCGCTCAAGCAGAGCCTCGTAGTGGCCGTCCAGGACCACGGCCAGCC  
CCCTCTCTCCGCCACTGTACGCTCACCGTGGCCGTGGCCGACAGCATCCCCAAGTCTGGCGGACCTCGGCAG  
CCTCGAGTCTCCAGCTAACTCTGAAACCTCAGACCTCACTCTGTACCTGGTGGTAGCGGTGGCCGCGGTCTCCTG  
CGTCTTCTGGCCTTCGTCTATCTTGTCTGGCGCTCAGGCTGCGGCGCTGGCACAAAGTACGCCCTGCTGCAGGC  
TTCAGGAGGCGGCTTGACAGGAGCGCCGGCGTCGCACTTTGTGGCGGTGGACGGGGTGCAGGCTTTCCTGCAGAC  
CTATTCCCACGAGGTTTCCCTCACCACGGACTCGCGGAAGAGTCACTGATCTTCCCCAGCCCACTATGCAGA  
CATGCTCGTCAGCCAGGAGAGCTTTGAAAAAGCGAGCCCCCTTTTGTCTGTCAGGTGATTGCGTATTTTCTAAAGA  
CAGTCATGGGTTAATTGAGGTGAGTTTATATCAAATCTTCTTTTCTTTTAAATTGCTCTGTCTCCCAAGC  
TGGAGTGACGCGGTACGATCATAGCTCACTGCGGCCTCAAACCTCCTAGGCTCAAGCAATTATCCCACCTTTGCT  
CCGGTGTAACAGGGACTACAGGTGCAAGCCACCTACTGTCTGCCTATCTATCTATCTATCTATCTATCTATCTAT  
CTATCTATCTATCTATCTATTACTTTCTTGTACAGACGGGAGTCTCACGCCTGTAATCCAGTACTTTGGGAGGC  
CGAGGCGGGTGGATCACCTGAGGTTGGGAGTTTGAGACCAGCTGACCAACATGGAGAAACCCGCTCTATACTAA  
AAAAATACAAAATTAGCCGGGCGTGGTGGTGCATGTCTGTAATCCAGCTACTTGGGAGGCTGAGTCAGGAGAAT  
TGCTTTAACCTGGGAGGTGGAGGTGCAATGAGCTGAGATTGTGCCATTGCACTCCAGCCTGGGCAACAAGAGTG  
AAACTCTATCTCA

79/615

**FIGURE 78**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48306
><subunit 1 of 1, 916 aa, 1 stop
><MW: 100204, pI: 4.92, NX(S/T): 4
MIPARLHRDYKGLVLLGILLGTLWETGCTQIRYSVPEELEKGSRVGDISRDIGLEPRELAERGVRIIPRGRTQLF
ALNPRSGSLVTAGRIDREELCMGAIKCQLNLDILMEDKVKIYGVEVEVRDINDNAPYFRESELEIKISENAATEM
RFPLPHAWDPDIGKNSLQSYELSPNTHFSLIVQNGADGSKYPELVLKRALDREEKAAHHLVLTASDGGDPVRTGT
ARIRVMVLDANDNAPAFAPQPEYRASVPENLALGTQLLVVNATDPDEGVNAEVRYSFYVDDKAAQVFKLDCNSGT
ISTIGELDHESGFGYQMEVQAMDNAGYSARAKVLITVLDVNDNAPEVVLTSLASSVPENSPRGTLIALNVNDQD
SEENGQVICFIQGNLPFKLEKSYGNYYSLVTDIVLDREQVPSYNITVTATDRGTPPLSTETHISLNVADTNDNPP
VFPQASYSAYIPENNPRGVSLVSVTAHDPDCEENAQITYSLAENTIQGASLSSYVSINSDTGVLYALSSFDYEQF
RDLQVKVMARDNGHPPLSSNVSLSLFVLDQNDNAPEILYPALPTDGSTGVELAPRSAEPGYLVTKVVAVDRDSGQ
NAWLSYRLLKASEPGLFSVGLHTGEVRTARALLDRDALKQSLVVAVQDHGQPPLSATVTTLTVAVADSIPOVLADL
GSLESPANSETSDLTLYLVVAVAAVSCVFLAFVILLALLALRLRRWHKSRLLQASGGGLTGAPASHFVGVDGVQAFI
QTYSHEVSLTTDSRKSHLIFPQPNYADMLVSQESFEKSEPLLLSGDSVFESKDSHGLIEVSLYQIFFLFFNCVS
QAGVQRYDHSSLRPQTPRLKQLSHLCLRCNRDYRCKPPTVCLSIYLSIYLSIYLSIYLLSCTDGS LTPVIPVLW
EAEAGGSPEVGS LRPA
```

**Signal sequence:**

amino acids 1-30

**Transmembrane domains:**

amino acids 693-711, 809-823, 869-888

80/615

**FIGURE 79**

AGCCGCTGCCCCGGGCGGGCGCCCGCGGCGGCACCA**ATG**AGTCCCCGCTCGTGCCCTGCGTTGCTGCGCCTCCTC  
GTCTTCGCCGTCTTCTCAGCCGCCGCGAGCAACTGGCTGTACCTGGCCAAGCTGTCTGTCGGTGGGGAGCATCTCA  
GAGGAGGAGACGTGCGAGAACTCAAGGGCCTGATCCAGAGGCAGGTGCAGATGTGCAAGCGGAACCTGGGAAGTC  
ATGGACTCGGTGCGCCGCGGTGCCAGCTGGCCATTGAGGAGTGCCAGTACCAGTTCCGGAACCGGCGCTGGAAC  
TGCTCCACACTCGACTCCTTGCCCGTCTTCGGCAAGGTGGTGACGCAAGGGACTCGGGAGGCGGCCTTCGTGTAC  
GCCATCTCTTCGGCAGGTGTGGCCTTTGCAGTGACGCGGGCGTGCAGCAGTGGGGAGCTGGAGAAGTGCGGCTGT  
GACAGGACAGTGTCATGGGGTCAAGCCACAGGGCTTCCAGTGGTCAGGATGCTCTGACAACATCGCCTACGGTGTG  
GCCTTCTCACAGTCGTTTGTGGATGTGCGGGAGAGAAGCAAGGGGGCCTCGTCCAGCAGAGCCCTCATGAACCTC  
CACAACAATGAGGCCCGGCAGGAAGGCCATCCTGACACACATGCGGGTGAATGCAAGTGCCACGGGGTGTGAGGC  
TCCTGTGAGGTAAAGACGTGCTGGCGAGCCGTGCCGCCCTTCCGCCAGGTGGGTACGCACTGAAGGAGAAGTTT  
GATGGTGCCACTGAGGTGGAGCCACGCCGCGTGGGCTCCTCCAGGGCACTGGTACCACGCAACGCACAGTTCAAG  
CCGCACACAGATGAGGACCTGGTGTACTTGAGCCTAGCCCCGACTTCTGTGAGCAGGACATGCGCAGCGGCGTG  
CTGGGCACGAGGGGCGCACATGCAACAAGACGTCCAAGGCCATCGACGGCTGTGAGCTGCTGTGCTGTGGCCGC  
GGCTTCCACACGCGCGCAGGTGGAGCTGGCTGAACGCTGCAGCTGCAAATCCACTGGTGTGCTTCGTCAAGTGC  
CGGCAGTGCCAGCGGCTCGTGGAGTTGCACACGTGCCGA**TGA**CCGCCCTGCCCTAGCCCTGCGCCGGCAACCACCTA  
GTGGCCCAGGGAAGGCCGATAATTTAAACAGTCTCCCACCACCTACCCCAAGAGATACTGGTTGTATTTTTTGT  
CTGGTTTGGTTTTTTGGGTCTCATGTTATTTATTGCCGAAACCAGGCAGGCAACCCCAAGGGCACCAACCAGGGC  
CTCCCCAAAGCCTGGGCCTTTGTGGCTGCCACTGACCAAGGGACCTTGCTCGTGCCGCTGGCTGCCCGCATGTG  
GCTGCCACTGACCACTCAGTTGTTATCTGTGTCCGTTTTTCTACTTGCAGACCTAAGGTGGAGTAACAAGGAGTA  
TTACCACCACATGGCTACTGACCGTGTCTCGGGGAAGAGGGGGCCTTATGGCAGGGAAAAATAGGTACCGACTTG  
ATGGAAGTCACACCCTCTGGAAAAAGAACTCTTAACCTCTCCAGCACACATACACATGGACTCCTGGCAGCTTGA  
GCCTAGAAGCCATGTCTCTCAAATGCCCTGAGAAAGGGAACAAGCAGATACCAGGTCAAGGGCACCAAGTTTCAAT  
TCAGCCCTTACATGGACAGCTAGAGGTTGATATCTGTGGGTCTTCCAGGCAAGAAGAGGGAGATGAGAGCAAG  
AGACGACTGAAGTCCCACCCTAGAACCCAGCCTGCCCCAGCCTGCCCCCTGGGAAGAGGAACTTAACCACTCCCC  
AGACCCACCTAGGCAGGCATATAGGCTGCCATCCTGGACCAGGGATCCCGGCTGTGCCTTTGCAGTCATGCCCGA  
GTACCTTTTACAGCGCTGTTCTCCATGAACTGAAAAACACACACACACACACACACACACACACACACACAC  
ACACACACACGGACACACACACACACACCTGCGAGAGAGAGGGAGGAAAGGGCTGTGCCTTTGCAGTCATGCCCGAG  
TCACCTTTTACAGCACTGTTCTCTC

81/615

**FIGURE 80**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48328  
<subunit 1 of 1, 351 aa, 1 stop  
<MW: 39052, pI: 8.97, NX(S/T): 2  
MSPRSCLRSLRLLVFAVFSAAAASNWLYLAKLSSVGSISEEETCEKLKGLIQRQVQMCKRNLEVMDSVRRGAQLAI  
EECQYQFRNRRWNCSTLDSLVPVFGKVVTQGTREAAFVYAISSAGVAFVTRACSSGELEKCGCDRTVHGVSPQGF  
QWSGCSNDIAYGVAFSQSFVDVRERSKGASSSRALMNLHNNEAGRKAILTHMRVECKCHGVSGSCEVKTWCRAVP  
PFRQVGHALKEKFDGATEVEPRRVGSSRALVPRNAQFKPHTDEDLVYLEPSPDFCEQDMRSGVLGTRGRTCNKTS  
KAIDGCELLCCGRGFHTAQVELAERCSCKFHWCCFVKCRQCQRLVELHTCR

**Important features:****Signal peptide:**

amino acids 1-22

**N-glycosylation sites.**

amino acids 88-91 and 297-300

**Wnt-1 family signature.**

amino acids 206-215

**Homologous region to Wnt-1 family proteins**

amino acids 183-235, 305-350, 97-138, 53-92 and 150 -174

82/615

**FIGURE 81**

CCGAGCCGGGCGCGCAGCGACGGAGCTGGGGCCGGCCTGGGACCATGGGCGTGAGTGCAATCTACGGATCAGTCT  
CTGATGGTGGGTCGTTAACCTCAGTGGGGACTCCAAGATTTCCATGAAGAAAATCAGTTGTCTTCATTCAAGAAT  
TGGGGTCTGGCTCAGAATTCCTGCAGCTGGTGAATACTGTTTTCTAGAAGAGGTTTAATTAATGCCTGCAGTCT  
GACATGTTCCCGATTTGAGGTGAAACCATGAAGAGAAAATAGAATACTTAATAATGCTTTTCCGCAACCGCTTCT  
TGCTGCTGCTGGCCCTGGCTGCGCTGCTGGCCTTTGTGAGCCTCAGCCTGCAGTTCTTCCACCTGATCCCGGTGT  
CGACTCCTAAGAATGGAATGAGTAGCAAGAGTCGAAAGAGAATCATGCCCGACCCTGTGACGGAGCCCCCTGTGA  
CAGACCCCGTTTATGAAGCTCTTTTGTACTGCAACATCCCCAGTGTGGCCGAGCGCAGCATGGAAGGTGATGCCC  
CGCATCATTTTAAGCTGGTCTCAGTGCATGTGTTCAATCGCCACGGAGACAGGTACCCACTGTATGTCATTCCCA  
AAACAAAGCGACAGCAAATGACTGCACTCTGGTGGCTAACAGGAAACCGTATCACCCAAAACCTGGAAGCTTTCA  
TTAGTCACATGTCAAAAGGATCCGGAGCCTCTTTCGAAAGCCCCCTTGAACCTCCTTGCCCTCTTTACCCAAATCACC  
CATGTGTGAGATGGGAGAGCTCACACAGACAGGAGTTGTGCAGCATTTGCAGAACGGTCAGCTGCTGAGGGATA  
TCTATCTAAAGAAACACAACTCCTGCCAATGATTGGTCTGCAGACCAGCTCTATTAGAGACCACTGGGAAAA  
GCCGGACCTACAAAGTGGGCTGGCCTTGCTTTATGGCTTTCTCCCAGATTTTGAAGTGAAGAAGATTTATTTCA  
GGCACCAGCCAAGTGGCTGTTCTGCTCTGGAAGCTGCTATTGCCCGTAAGAAACAGTATCTGGAAAAGGAGC  
AGCGTCGTGAGTACCTCCTACGTTTGAAGAACAGCCAGCTGGAGAAGACCTACGGGGAGATGGCCAAGATCGTGG  
ATGTCCCCACCAAGCAGCTTAGAGCTGCCAACCCCATAGACTCCATGCTCTGCCACTTCTGCCACAATGTGAGCT  
TTCCCTGTACCAGAAATGGCTGTGTTGACATGGAGCACTTCAAGGTAATTAAGACCCATCAGATCGAGGATGAAA  
GGGAAAGACGGGAGAAGAAATTGACTTCCGGTATTCTCTCCTGGGTGCCACCCCATCCTGAACCAACCATCG  
GCCGGATGCAGCGTGCCACCGAGGGCAGGAAAGAAGAGCTCTTGGCCCTCTACTCTGCTCATGATGTCACTCTGT  
CACCAGTTCTCAGTGCCTTGGGCCTTTCAGAAGCCAGGTTCCCAAGGTTTGCAGCCAGGTTGATCTTTGAGCTTT  
GGCAAGACAGAGAAAAGCCAGTGAACATTCCGTCCGGATTCTTTACAATGGCGTCGATGTACATTCCACACCT  
CTTCTGCCAAGACCACCACAAGCGTTCTCCCAAGCCCATGTGCCCGCTTGAAAACCTTGGTCCGCTTTGTGAAAA  
GGGACATGTTTGTAGCCCTGGGTGGCAGTGGTACAAATTATTATGATGCATGTACAGGGAAGGATTCATAAAGG  
TATGCAGTACAGCAGTATAGAATCCATGCCAATACAGAGCATAGGGAAAGGTCCACTTCTAGTTTTGTCTGTTAC  
TAAGGGTAGAAGATTATTGCTTTTTAAAGGCTAAATATTGTTTGTGGGAACACAGATGGTTGGGGTTGAACAGT  
AAGCACATTGCTGCAATGTGGTACGTGAATTGCTTGGTACAAAATGGCCAGTTCACAGAGGAATAGAAGGTACTT  
TATCATAGCCAGACTTCGCTTAGAATGCCAGAATAATATAGTTCAAGACCTGAAGTTGCCAATCCAAGTTTGCAC  
TCTTCTGGCCTGCCCCATGTTACTATGTGATGGAACAGCACACCTCAACCAAAATTTTTTTAATCTTAGACATT  
TTTACCTTGTCCTTGTTAAGAATTTCTTGAAGTGATTATCTAAAATAAAGGTTGGCAAACCTTTTTCTGTAAAGG  
GCCAGATTGTAAATATTTTCAAGCTGTGTGGACCAAAAGGCCACATACAGTCTCTGTGCATAACTACTCAACTCTGT  
TTCTGAAGCAGGAAAGCCACCACAGACAGTACATAAAGGAATATGTGTAGCTGGGTTCCAGGCCAGACAAAACA  
GATGGTGACCAGACTTGGCCCTGGGCTGTAGTTTGTGACCCCTCATCTAAAAAATAGGCTATACTACAATTGC  
ACTTCCAGCACTTTGAGAACGAGTTGAATACCAAGAATTATCAATGGTTCTCCAGTAACTTCTGCTAGAAAACA  
CAGAATTTGGTCTGTATCTGACACTAGAACAAAACCTTGAGGGTAAATAAACATTGAATTAGAATGAATCATAGAA  
AACTGATTAGAAGAATACTTGATGTTTATGATGATTGTTGTTACAAGATAGTTTTAAGTATGTTCTAAATATTTGT  
CTGCTGTAGTCTATTTGCTGTATATGCTGAAATTTTGTATGCCATTTAGTATTTTATAGTTTAGGAAAATATT  
TTCTAAGACCAGTTTTAGATGACTCTTATTCCTGTAGTAATATTCAATTTGCTGTACCTGCTTGGTGGTTAGAAG  
GAGGCTAGAAGATGAATTCAGGCACCTTCTTCCAATAAACTAATTATGGCTCATTCCCTTTGACAAGCTGTAGA  
ACTGGATTCATTTTTAAACCATTTTTCATCAGTTTCAAATGGTAAATTCTGATTGATTTTTAAATGCGTTTTTGA  
AGAATTTGCTATTAGGTAGTTTACAGATCTTTATAAGGTGTTTTATATATTAGAAGCAATTATAATTACATCTG  
TGATTTCTGAACATAATGGTGCTAATTACAGAGAAATGGAAGTGAAAGTGAGATTCTCTGTTGTGATCGGCATTCC  
AACTTTTTCTCTTTGTTTTGTCCAGTGTTGCATTTGAATATGTCTGTTTCTATAAATAAATTTTTTAAGAATAA

83/615

**FIGURE 82**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA48329
><subunit 1 of 1, 480 aa, 1 stop
><MW: 55240, pI: 9.30, NX(S/T): 2
MLFRNRFLLLLAIAALLAFVSLSLQFFHLIPVSTPKNGMSSKSRKRIMPDVTEPPVTDPVYEALLYCNIPSAE
RSMEGHAPHHFKLVSVHVFIRHGDRYPLYVIPKTKRPEIDCTLVANRKPYPHPKLEAFISHMSKSGSGASFESPLNS
LPLYPNHPLCEMGELTQTGVVQHLQNGQLLRDIYLLKKHKLLPNDWSADQLYLETTGKSRTLQSGLALLYGFLPDF
DWKKIYFRHQPSALFCSGSCYCPVRNQYLEKEQRRQYLLRLKNSQLEKTYGEMAKIVDVPTKQLRAANPIDSMC
HFCHNVSFPCTRNGCVDMEHFKVIKTHQIEDERERREKKLYFGYSLLGAHPILNQITIGRMQRATEGRKEELFALY
SAHDVTLSFVLSALGLSEARFPRFAARLIFELWQDREKPSSEHSVRILYNGVDVTFHTSFCQDHHKRSPKPMCPLE
NLVRFVVKRDMFVALGGSGTNYDACHREGF
```

**Signal sequence:**  
amino acids 1-18

84/615

**FIGURE 83**

TCTCGCAGATAGTAAATAATCTCGGAAAGGCGAGAAAGAAGCTGTCTCCATCTTGTCTGTATCCGCTGCTCTTGT  
GACGTTGTGGAGATGGGGAGCGTCCTGGGGCTGTGCTCCATGGCGAGCTGGATACCATGTTTGTGTGGAAGTGCC  
CCGTGTTTGCTATGCCGATGCTGTCTAGTGGAACAACCTCCACTGTAAGTAGATTGATCTATGCACCTTTCTTG  
CTGTGTGGAGTATGTGTAGCTTGTGTAATGTTGATACCAGGAATGGAAGAACAACGAATAAGATTCTTGGATTT  
TGTGAGAATGAGAAAGGTGTTGTCCCTTGTAAACATTTTGGTTGGCTATAAAGCTGTATATCGTTTTGTGCTTTGGT  
TTGGCTATGTTCTATCTTCTTCTCTCTTTACTAATGATCAAAGTGAAGAGTAGCAGTGATCCTAGAGCTGCAGTG  
CACATGGATTTTGGTTCTTTAAATTTGCTGCAGCAATTGCAATTATTATTGGGGCATTCTTCATTCCAGAAGGA  
ACTTTTACAACCTGTGTGGTTTTATGTAGGCATGGCAGGTGCCTTTTGTTCATCCTCATACAACCTAGTCTTACTT  
ATTGATTTTGCACATTCATGGAATGAATCGTGGGTTGAAAAAATGGAAGAAGGGAACCTCGAGATGTTGGTATGCA  
GCCCTGTATCAGCTACAGCTCTGAATTATCTGCTGTCTTTAGTTGCTATCGTCTGTTCTTTGTCTACTACACT  
CATCCAGCCAGTTGTTTCAAAAAACAAGGCGTTCATCAGTGTCAACATGCTCCTCTGCGTTGGTGCTTCTGTAAATG  
TCTATACTGCCAAAAATCCAAGAATCACAACCAAGATCTGGTTTGTACAGTCTTCAGTAATTACAGTCTACACA  
ATGATTTTGACATGGTCAGCTATGACCAATGAACCAGAAACAATTCGAACCCAAGTCTACTAAGCATAAATGGC  
TACAATACAACAAGCACTGTCCCAAGGAAGGGCAGTCAGTCCAGTGGTGGCATGCTCAAGGAATTATAGGACTA  
ATTCTCTTTTGTGTTGTGTATTTTATCCAGCATCCGTACTTCAAACAATAGTCAGGTTAATAAACTGACTCTA  
ACAAGTGATGAATCTACATTAATAGAAGATGGTGGAGCTAGAAGTGATGGATCACTGGAGGATGGGGACGATGTT  
CACCGAGCTGTAGATAATGAAAGGGATGGTGTCACTTACAGTTATTCCTTCTTTCACTTCATGCTTTTCTGGCT  
TCACCTTTATATCATGATGACCTTACCAACTGGTCCAGGTATGAACCTCTCGTGAGATGAAAAGTCAGTGGACA  
GCTGTCTGGGTGAAAATCTCTTCCAGTTGGATTGGCATCGTGTCTGTATGTTGGACACTCGTGGCACCAGTTGTT  
CTTACAAATCGTGATTTTGAAGTGAAGTGTGATGTTTTTGTCTCCATGTAACCTTCTCCAGTGTCTGGCATGAATTA  
GATTTTACTGCTTGTCTATTTTGTATTTTCTTACCAAGTGCATTGATATGTGAAGTAGAATGAATTCAGAGGAA  
AGTTTTATGAATATGGTGATGAGTTAGTAAAGTGGCCATTATTGGGCTTATCTCTGCTCTATAGTTGTGAAAT  
GAAGAGTAAAAACAATTTGTTTGAATTTTAAATATATTAGACCTTAAGCTGTTTTCAGCAAGCATTAAAGC  
AAATGTATGGCTGCCTTTTGAATATTTGATGTGTGCTGGCAGGATACTGCAAGAAGCAATGGTTTTATTTTAAA  
ATTTATAACAAGTCACTTAAATGCCAGTTGTCTGAAAAAATCTTATAAGGTTTACCCTTGATACGGAATTTACA  
CAGGTAGGGAGTGTTTAGTGGACAATAGTGTAGGTTATGGATGGAGGTGTCGGTACTAAATTGAATAACGAGTAA  
ATAATCTTACTTGGGTAGAGATGGCCTTTGCCAACAAAGTGAAGTGTGTTTGGTTGTTTTAACTCATGAAGTATG  
GGTTCAGTGGAATGTTTGGAACTCTGAAGGATTTAGACAAGGTTTTGAAAAGGATAATCATGGGTAGAAAGGAA  
GTGTTTTGAAAGTCACTTTGAAAGTTAGTTTTGGGCCAGCACGGTAGCTCACCTTGGTAATCCAGCACTTTG  
GGAGCTTAAGTGGGTAGATTACTTGAAGCCAGGAATTCAGACCAGCTTGGCACATGGTGAACCTGTTCTATAAAA  
ATAATCTGGCTTTGAGCATATGCTGTGGTCCAGCACTGAGAGGCTAGTGAAGATTGCTGAGCCAGAGCCAAAG  
GTTGCAGTGAGCAAGTCACGTCACTGCACCTAGCTGGCACAGAGTAAGCCAAAAAATATATATATATGAAAT  
CAAGGAGGCAAAATTTTGACAGGGAAGGAAGTAAGTGCAAAACCACTAGGCTTTAGTAGGTACTTATATAAAATC  
TAGTCCAGTTCTCTCATTTAAAAAATGAAGACACTGAAATACAGACTTAAATAGCTCAGATAGCTAATTAGGAA  
ATTTCAAGTTGGCCAATAATAGCATTCTCTGACATTTAAAAATAATTTCTATTCAAAATACATGCATATTGAT  
TTACACCTCATACTGTGATAATTAATGTGATGTGGATTGCTGGTGTCCAGCATGACCCATAAACAGGTCCAGAAGA  
ATGATGGAATGTTTGAATAAACTCCTGCTTATAGTATACTACACAGTTCAAAAGATGTTTAAATGCTTTTGT  
ATTTACTGCCATGTAATTGAAATATATAGATTATTGTAACCTTTCAACCTGAAAATCAAGCAGTATGAGAGTTTA  
GTTATTTGTATGTGTCACTAGTGTCTAATGAAGCTTTTAAATCTACAATTTCTTCTTTAAAAATATTTTAAAT  
GTGAATGGAATATAACAATTCAGCTTAATCCCCAACCTTATTCTGTGTGTAGACATTGTATCCACAATTTTGA  
ATGGCTGTGTTTTACCTCTAAATAAATGAATTCAGAGAAAAA

85/615

**FIGURE 84**

MGSVLGLCSMASWIPCLCGSAPCLLCRCCPSGNNSTVTRLIYALFLLVGVCVACVMLIPGMEEQLNKIPGFCENE  
KGVVPCNILVGYKAVYRLCFGLAMFYLLSLLMIKVKSSSDPRAAVHNGFWFFKFAAAIAIIIGAFFIPEGTFTT  
VWFYVGMAGAFCFILIQVLVLLIDFAHSWNEVVEKMEEGNSRCWYAALLSATALNYLLSLVAIVLFFVYYTHPAS  
CSENKAFISVNMLLCVGASVMSILPKIQESQPRSGLLQSSVITVYTMYLTSAMTNEPETNCNPSLLSIIGYNTT  
STVPKEGQSVQWWHAQGIIGLILFLLCVFYSSIRTSNNSQVNKLTLTSDESTLIEDGGARSDGSLEDGDDVHRAV  
DNERDGVITYSYFFHEMLFLASLYIMMTLTNWSRYEPSREMKSQWTAVVWKISSSWIGIVLYVWTLVAPLVLTNRDFD



86/615

**FIGURE 85**

AACAAAGTTTCAGTGACTGAGAGGGCTGAGCGGAGGCTGCTGAAGGGGAGAAAGGAGTGAGGAGCTGCTGGGCAGA  
GAGGGACTGTCCGGCTCCCAGATGCTGGGCCTCCTGGGGAGCACAGCCCTCGTGGGATGGATCACAGGTGCTGCT  
GTGGCGGTCTGCTGCTGCTGCTGCTGCTGGCCACCTGCCTTTTCCACGGACGGCAGGACTGTGACGTGGAGAGG  
AACCGTACAGCTGCAGGGGGAAACCGAGTCCGCCGGGCCAGCCTTGGCCCTTCCGGCGCGGGGCCACCTGGGA  
ATCTTTCACCATCACCGTCATCCTGGCCACGTATCTCATGTGCCGAATGTGGGCCTCCACCACCACCACCC  
CGCCACACCCCTCACCACCTCCACCACCACCACCCCCACCGCCACCATCCCCGCCACGCTCGCTGAGGGCTGC  
TGTCGCCGGTGCCTGTGGACAGCAGCTGCCCTGCCCTCCCATCTGTTCCAGGACAAGTGGACCCCATGTTTCC  
ATGTGGAAGGATGCATCTCTGGGGTGAACGAGGGGAACAATAGACTGGGGCTTGCTCCAGCTGCATTTGCATGGC  
ATGCCCCAGTGTACTATGGCAGCAGAGAATGGAGGAACACTGGGTCTGCAGTGCTGAAGGGTTTGGGGAGTGGAG  
AGCAAGGGTGCTCTTTCGGGGCTGGACAGCCCGTCTTGTGACAGTGAAGTGGAGTGGAGTGGAGTGGAGTGGAG  
GTGTCTTGGCAGAGCCAGCACACAAGTGGATGTGAAGTGCCCGTCTTGACCTCCTCATCAGGCTGCTGCAGGCCT  
CTGGCGGGCAGGGCACTGGGAGAGGCCCTGAGAATGTCCTTTTGGTTTGGAGAAGGCAGTGTGAGGCTGCACAGT  
CAATTCATCGGTGCCTTAGTCCAAGAAAATAAAAACCACTAAGAAGCTTTAAAAAAAAAAAAAAAAAAAAA

87/615

**FIGURE 86**

MLGLLGSTALVGWITGA AVAVLLLLLLLATCLFHGRQDCDVERNRTAAGGNRVRAQWPWFRRRGHLGIFHHHRH  
PGHVSHVFPNVGLHHHHHPRHTPHHLHHHHHPRHHPRHAR

88/615

**FIGURE 87**

CCCACGCGTCCGTCCTAGTCCCCGGGCCAACTCGGACAGTTTGCTCATTATTATGCAACGGTCAAGGCTGGCTTGT  
GCCAGAACGGCGCGCGCGCGCGCACGCACGCACACACGCGGGGGAACTTTTTTAAAAATGAAAGGCTAGAAGA  
GCTCAGCGGCGGGCGCGGGCGCTGCGCGAGGGCTCCGGAGCTGACTCGCCGAGGCAGGAAATCCCTCCGGTCGCGA  
CGCCCCGGCCCCGGCTCGGCGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGTGTCT  
GCCGGCGACGATGGCAGCGCGCCCGCTGCCCGTGTCCCCCGCCCGCGCCCTCCTGCTCGCCCTGGCCGGTGTCTCT  
GCTCGCGCCCTGCGAGGCCCGAGGGGTGAGCTTATGGAACCAAGGAAGAGCTGATGAAGTTGTGAGTGCCTCTGT  
TCGGAGTGGGGACCTCTGGATCCAGTGAAGAGCTTCGACTCCAAGAATCATCCAGAAGTGTGAATATTCGACT  
ACAACGGGAAAGCAAAGAAGTATCATATAATCTGGAAGAAATGAAGGTCTCATTGCCAGCAGTTTCACGGAAAC  
CCACTATCTGCAAGACGGTACTGATGTCTCCCTCGCTCGAAATTACACGGGTCACTGTTACTACCATGGACATGT  
ACGGGGATATTCTGATTACGAGTCAGTCTCAGCACGTGTTCTGGTCTCAGGGGACTTATTGTGTTTGAAAATGA  
AAGCTATGTCTTAGAACC AATGAAAAGTGCACCAACAGATACAACTCTTCCAGCGAAGAAGCTGAAAAGCGT  
CCGGGGATCATGTGGATCACATCACAACACACCAACCTCGCTGCAAAGAATGTGTTTCCACCACCTCTCAGAC  
ATGGGCAAGAAGGCATAAAAGAGAGACCCTCAAGGCAACTAAGTATGTGGAGCTGGTGTGCTGGCAGACAACCG  
AGAGTTTCAGAGGCAAGGAAAAGATCTGGAAGAACTTAAGCAGCGATTAATAGAGATTGCTAATCACGTTGACAA  
GTTTTACAGACCACTGAACATTCGGATCGTGTGGTAGGCGTGGAAGTGTGGAATGACATGGACAAATGCTCTGT  
AAGTCAGGACCCATTACACAGCCTCCATGAATTTCTGGACTGGAGGAAGATGAAGCTTCTACCTCGCAAATCCCA  
TGACAATGCGCAGCTTGTGAGTGGGGTTTATTTCCAAGGACCACCATCGGCATGGCCCCAATCATGAGCATGTG  
CACGGCAGACCACTGCGGGGAATTGTGATGGACCATTCAGACAATCCCTTGGTGCAGCCGTGACCCCTGGCACA  
TGAGCTGGGCCACAATTTGGGATGAATCATGACACACTGGACAGGGGCTGTAGCTGTCAAATGGCGGTTGAGAA  
AGGAGGCTGCATCATGAACGCTTCCACCGGTACCCATTTCCCATGGTGTTCAGCAGTTCAGCAGGAAGGACTT  
GGAGACCAGCCTGGAGAAAGGAATGGGGGTGTGCCTGTTTTAACCTGCCGGAAGTCAGGGAGTCTTTGGGGGCCA  
GAAGTGTGGGAACAGATTGTGGAAGAAGGAGAGAGTGTGACTGTGGGGAGCCAGAGGAATGTATGAATCGCTG  
CTGCAATGCCACCACCTGTACCCTGAAGCCGACGCTGTGTGCGCACATGGGCTGTGCTGTGAAGACTGCCAGCT  
GAAGCCTGCAGGAACAGCGTGCAGGGACTCCAGCAACTCCTGTGACCTCCAGAGTTCTGCACAGGGGCCAGCCC  
TCACTGCCCAGCCAATGTGTACCTGCACGATGGGCACTCATGTGAGGATGTGGACGGCTACTGCTACAATGGCAT  
CTGCCAGACTCACGAGCAGCAGTGTGTGACGCTCTGGGGACCAGGTGCTAAACCTGCCCTGGGATCTGCTTTGA  
GAGAGTCAATCTGCAGGTGATCCTTATGGCAACTGTGGCAAAGTCTCGAAGAGTTCTTTTGCCAAATGCGAGAT  
GAGAGATGCTAAATGTGGAAGAAATCCAGTGTCAAGGAGGTGCCAGCCGGCCAGTCATTGGTACCAATGCCGTTTC  
CATAGAAACAAACATCCCTCTGCAGCAAGGAGGCCGATTCTGTGCCGGGGACCCACGTGTACTTGGGCGATGA  
CATGCCGGACCCAGGGCTTGTGCTTGCAGGCACAAAGTGTGCAGATGGAAAAATCTGCCTGAATCGTCAATGTCA  
AAATATTAGTGTCTTTGGGGTTCACGAGTGTGCAATGCAGTGCCACGGCAGAGGGGTGTGCAACAACAGGAAGAA  
CTGCCACTGCGAGGCCCACTGGGCACCTCCCTTCTGTGACAAGTTTGGCTTTGGAGGAAGCACAGACAGCGGCC  
CATCCGGCAAGCAGAAAGCAAGGCAGGAAGCTGCAGAGTCCAACAGGGAGCGCGGCCAGGGCCAGGAGCCCGTGGG  
ATCGCAGGAGCATGCGTCTACTGCCTCACTGACACTCATCTGAGCCCTCCCATGACATGGAGACCGTGACCAGTG  
CTGCTGCAGAGGAGTTCACGCGTCCCCAAGGCCTCCTGTGACTGGCAGCATTGACTCTGTGGCTTTGCCATCGTT  
TCCATGACAACAGACACAACACAGTTCTCGGGGCTCAGGAGGGGAAGTCCAGCCTACCAGGCAGTCTGCAGAAA  
CAGTGCAGGAAGGGCAGCGACTTCTGTTGAGCTTCTGCTAAAACATGGACATGCTTCAGTGTCTCTCTGAG  
AGAGTAGCAGGTTACCACTCTGGCAGGCCCCAGCCCTGCAGCAAGGAGGAAGAGGACTCAAAAGTCTGGCCTTTC  
ACTGAGCCTCCACAGCAGTGGGGGAGAAGCAAGGGTTGGGCCAGTGTCCCCTTTCCCACTGACACCTCAGCCT  
TGGCAGCCCTGATGACTGGTCTCTGGCTGCAACTTAATGCTCTGATATGGCTTTTAGCATTTATATATGAAAAT  
AGCAGGGTTTTAGTTTTTAATTTATCAGAGACCCTGCCACCATTCCATCTCCATCCAAGCAAACTGAATGGCAA  
TGAACAAACTGGAGAAGAAGGTAGGAGAAAGGGCGGTGAACCTTGGCTCTTTGCTGTGGACATGCGTGACCAGC  
AGTACTCAGGTTTGAGGGTTTGAGAAAGCCAGGGAACCCACAGAGTCACCAACCCTTCATTTAACAAGTAAGAA  
TGTTAAAAAGTGAAAACAATGTAAGAGCCTAACTCCATCCCCCGTGGCCATTACTGCATAAAATAGAGTGCATTT  
GAAAT

89/615

**FIGURE 88**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49624
><subunit 1 of 1, 735 aa, 1 stop
><MW: 80177, pI: 7.08, NX(S/T): 5
MAARPLPVSPARALLALAGALLAPCEARGVSLWNQGRADEVVVSASVRSGDLWIPVKSFD SKNHPEVLNIRLQRE
SKELIINLERNEGLIASSFTETHYLQDGTDVSLARNYTGHCYYHGHVVRGYSDSAVSLSTCSGLRGLIVFENESYV
LEPMKSATNRYKLFPAAKKLSVRGSCGSHHNTPNLAAKNVFPPPSQTWARRHKRETILKATKYVELVIVADNREFQ
RQGKDLEKVKQRLIEIANHVDKFYRPLNIRIVLVGVEVWVNDMDKCSVSQDPFTSLHEFLDWRKMKLLPRKSHDNA
QLVSGVYFQGTITIGMAPIMSMCTADQSGGIVMDHSDNPLGAAVTLAHELGHNF GMNHDTLDRGCSCQMAVEKGGC
IMNASTGYPPFMVFSSCSRKDLETSLEKGMGVCLFNLPEVRESFGGQKCGNRFEVEEGEECDCEPEECMNRCCNA
TTCTLKPDCAVCAHGLCCEDCQLKPAGTACRDSSNSCDLPEFCTGASPHCPANVYLHDGHSCQDQVDGYCYNGICQT
HEQQCVTLWGPGAKPAPGICFERVNSAGDPYGNCGKVS KSSFAKCEMRDAKCGKIQCQGGASRPVIGTNAVSIET
NIPLOQGGRI LCRGTHVYLGDDMPDPGLVLAGTKCADGKICLN RQCQNISVFGVHECAMQCHGRGVCNNRKNCHC
EAHWAPPFCDFKFGFGSTDSGPIRQAEARQEAESNRERGGQGE PVGSQEHASTASLTLLI
```

**Signal peptide:**  
amino acids 1-28

90/615

**FIGURE 89**

CTGCTGCATCCGGGTGTCTGGAGGCTGTGGCCGTTTTGTTTTCTTGGCTAAAAATCGGGGGAGTGAGGCGGGCCGG  
CGCGGCGCGACACCGGGCTCCGGAACCACTGCACGACGGGGCTGGACTGACCTGAAAAAAATGCTCTGGATTTCTA  
GAGGGCTTGAGATGCTCAGAATGCATTGACTGGGGGGAAAAGCGCAATACTATTGCTTCCATTGCTGCTGGTGTA  
CTATTTTTTACAGGCTGGTGGATTATCATAGATGCAGCTGTTATTTATCCCACCATGAAAGATTTCAACCACTCA  
TACCATGCCTGTGGTGTATAGCAACCATAGCCTTCCTAATGATTAATGCAGTATCGAATGGACAAGTCCGAGGT  
GATAGTTACAGTGAAGGTTGTCTGGGTCAAACAGGTGCTCGCATTGGCTTTTCGTTGGTTTCATGTTGGCCTTT  
GGATCTCTGATTGCATCTATGTGGATTCTTTTTGGAGGTTATGTTGCTAAAGAAAAAGACATAGTATACCCCTGGA  
ATTGCTGTATTTTTCCAGAATGCCTTCATCTTTTTTGGAGGGCTGGTTTTTAAGTTTGGCCGCACTGAAGACTTA  
TGGCAGTGAACACATCTGATTTCCACAGCACAAACAGCCCTGCATGGGTTTGTGTTTTTTTTACTGCTCACTCC  
CAACCTTTTGTAATGCCATTTTCTAACTTATTTCTGAGTGTAGTCTCAGCTTAAAGTTGTGTAATACTAAAATC  
ACGAGAACACCTAAACAACAACCAAAAAATCTATTGTGGTATGCACCTTGATTAACCTATAAAATGTTAGAGGAAAC  
TTTCACATGAATAATTTTTGTCAAATTTTATCATGGTATAATTTGTAAAAATAAAAAGAAATTACAAAAGAAATT  
ATGGATTTGTCAATGTAAGTATTTGTCAATCTGAGGTCCAAAACCACAATGAAAGTGCTCTGAAGATTTAATGT  
GTTTATTCAAATGTGGTCTCTTCTGTGTCAAATGTTAAATGAAATATAAACATTTTTTAGTTTTTAAATATTCC  
GTGGTCAAATTTCTTCCTCACTATAATTTGGTATTTACTTTTACCAAAAATTCTGTGAACATGTAATGTAAGTGGC  
TTTTGAGGGTCTCCCAAGGGGTGAGTGGACGTGTTGGAAGAGAGAAGCACCATGGTCCAGCCACCAGGCTCCCTG  
TGTCCCTTCCATGGGAAGGTCTCCGCTGTGCCTCTCATTTCCAAGGGCAGGAAGATGTGACTCAGCCATGACACG  
TGGTTCTGGTGGGATGCACAGTCACTCCACATCCACCACTG

91/615

**FIGURE 90**

MSGFLEGLRCSECIDWGEKRNTIASIAAGVLFFTGWIIIDA AVIYPTMKDFNHSYHACGVIATIAFLMINAVSN  
GQVRGDSYSEGCLGQTGARIWLFVGFMLAFGSLIASMWILFGGYVAKEKDIVYPGIAVEFFQNAFIFFGGLVFKFG  
RTEDLWQ

92/615

**FIGURE 91**

CGACGCCGGCGT**GATG**TGGCTTCCGCTGGTGCTGCTCCTGGCTGTGCTGCTGCTGGCCGTCTCTGCAAAGTTTA  
CTTGGGACTATTCTCTGGCAGCTCCCCGAATCCTTTCTCCGAAGATGTCAAACGGCCCCCAGCGCCCTGGTAAC  
TGACAAGGAGGCCAGGAAGAAGGTTCTCAAACAAGCTTTTTCAGCCCAACCAAGTGCCGGAGAAGCTGGATGTGGT  
GGTAATTGGCAGTGGCTTTTGGGGGCTGGCTGCAGCTGCAATTCTAGCTAAAGCTGGCAAGCGAGTCTGGTGCT  
GGAACAACATACCAAGGCAGGGGGCTGCTGTACATACCTTTGGAAAGAATGGCCTTGAATTTGACACAGGAATCCA  
TTACATTGGGCGTATGGAAGAGGGGCAGCATTTGGCCGTTTTATCTTGGACCAGATCACTGAAGGGCAGCTGGACTG  
GGCTCCCCGTGCTCTCCTTTTGACATCATGGTACTGGAAGGGCCCAATGGCCGAAAGGAGTACCCCATGTACAG  
TGGAGAGAAAGCCTACATTCAGGGCCTCAAGGAGAAGTTTCCACAGGAGGAAGCTATCATTTGACAAGTATATAAA  
GCTGGTTAAGGTGGTATCCAGTGGAGCCCCCTCATGCCATCCTGTTGAAATTCCTCCCATTTGCCCGTGGTTTCAGCT  
CCTCGACAGGTGTGGGCTGCTGACTCGTTTTCTCTCCATTCTTCAAGCATCCACCAGAGCCTGGCTGAGGTCTCT  
GCAGCAGCTGGGGGCTCCTCTGAGCTCCAGGCAGTACTCAGCTACATCTTCCCCACTTACGGTGTACCCCCAA  
CCACAGTGCCTTTTCCATGCACGCCCTGCTGGTCAACCACTACATGAAAGGAGGCTTTTATCCCCGAGGGGGTTC  
CAGTGAAATTCCTTCCACACCATCCCTGTGATTGAGCGGGCTGGGGGCGCTGTCTCTACAAAGGGCCACTGTGCA  
GAGTGTGTGCTGGACTCAGCTGGGAAAGCCTGTGGTGTGAGTGTGAAGAAGGGGCATGAGCTGGTGAACATCTA  
TTGCCCATCGTGGTCTCCAACGCAGGACTGTTCAACACCTATGAACACCTACTGCCGGGGAACGCCGCTGCCT  
GCCAGGTGTGAAGCAGCAACTGGGGACGGTGCGGGCCCGCTTAGGCATGACCTCTGTTTTCTCTGCTGCGAGG  
CACCAAGGAAGACCTGCATCTGCCGTCCACCACTACTATGTTTACTATGACACGGACATGGACCAGGCGATGGA  
GCGCTACGTCTCCATGCCAGGGAGAGGCTGCGGAACACATCCCTCTTCTTCTTCTCGCTTTCCCATCAGCCAA  
AGATCCGACCTGGGAGGACCGATTCCAGGCCGGTCCACCATGATCATGCTCATACCCACTGCCTACGAGTGGTT  
TGAGGAGTGGCAGGCGGAGCTGAAGGGAAAGCGGGGCAGTGAATATGAGACCTTCAAAAACCTCTTTTGTGGAAGC  
CTCTATGTGCTAGTGGTCTGAAACTGTTCCACAGCTGGAGGGGAAGGTGGAGAGTGTGACTGCAGGATCCCCACT  
CACCAACAGTTCTATCTGGCTGCTCCCCGAGGTGCTGCTACGGGGCTGACCATGACCTGGGCCGCTGCACCC  
TTGTGTGATGGCCTCCTTGAGGGCCAGAGCCCCATCCCCAACCTCTATCTGACAGGCCAGGATATCTTCACCTG  
TGACTGGTGGGGGCCCTGCAAGGTGCCCTGCTGTGCAGCAGCGCCATCCTGAAGCGGAACCTGTACTCAGACCT  
TAAGAATCTTGATTCTAGGATCCGGGCACAGAAGAAAAAGAATT**AG**TTCCATCAGGGAGGAGTCAAGGAATTTG  
CCCAATGGCTGGGGCATCTCCCTTGACTTACCCATAATGTCTTTCTGCATTAGTTCCTTGCACGTATAAAGCACT  
CTAATTTGGTTCTGATGCCTGAAGAGAGGGCTAGTTTAAATCACAATTCCGAATCTGGGGCAATGGAATCACTGC  
TTCCAGCTGGGGCAGGTGAGATCTTTACGCCTTTTATAACATGCCATCCCTACTAATAGGATATTGACTTGGATA  
GCTTGATGTCTCATGACGAGCGGCGCTCTGCATCCCTCACCCATGCCTCCTAACTCAGTGATCAAAGCGAATATT  
CCATCTGTGGATAGAACCCCTGGCAGTGTGTGCTCAGCTCAACCTGGTGGGTTCACTTCTGTCTTGGGCTTCTGCT  
CTCATTCAATTTAGTGCTACGCTGCACAGTTCTACACTGTCAAGGGAAAAGGGAGACTAATGAGGCTTAACCTCAA  
ACCTGGGCGTGGTTTTGGTTGCCATTCCATAGGTTTGGAGAGCTCTAGATCTCTTTTGTGCTGGGTTCACTGGCT  
CTTCAGGGGACAGGAAATGCCTGTGTCTGGCCAGTGTGGTTCTGGAGCTTTGGGGTAACAGCAGGATCCATCAGT  
TAGTAGGGTGCATGTGATGATCATATCCAATTCATATGGAAGTCCCGGTCTGTCTTCTTATCATCGGGGTG  
GCAGCTGGTTCTCAATGTGCCAGCAGGGACTCAGTACCTGAGCCTCAATCAAGCCTTATCCACCAATACACAGG  
GAAGGGTGTGACAGGGAAGGGTGACATCAGGAGTCAGGGCATGGACTGGTAAGATGAATACTTTGCTGGGCTGAA  
GCAGGCTGCAGGGCATTCCAGCCAAGGGCACAGCAGGGGACAGTGCAGGGAGGTGTGGGGTAAGGGAGGGAAAGTC  
ACATCAGAAAAGGGAAAGCCACGGAATGTGTGTGAAGCCCAGAAATGGCATTGTGAGTTAATTAGCACATGTGAG  
GGTTAGACAGGTAGGTGAATGCAAGCTCAAGGTTTGGAAAAATGACTTTTCAGTTATGTCTTTGGTATCAGACAT  
ACGAAAGGTCTCTTTGTAGTTCGTGTTAATGTAACATTAATAAATTTATTGATTCCATTGCTTTAAAAA

93/615

**FIGURE 92**

MWLPLVLLLAVALLLAVLCKVYLGLFSGSSPNPFSEDKRPPAPLVTDKEARKKVLKQAFSANQVPEKLDVVVIGS  
GFGGLAAAAAILAKAGKRVLVLEQHTKAGGCCHTFGKNGLEFDTGIHYIGRMEEGSIGRFILDQITEGQLDWAPLS  
SPFDIMVLEGPNGRKEYPMYSGEKAYIQGLKEKFPQEEAIIIDKYIKLVKVVSSGAPHAILLKFLPLPVVQLLDRC  
GLLTRFSPFLQASTQSLAEVLQQLGASSELQAVLSYIFPTYGVTPNHSAFSMHALLVNHMKGGFYPRGGSSEIA  
FHTIPVIQRAGGAVLTATVQSVLLDSAGKACGVSVKKGHELVNIYCPIVVSNAGLFNTYEHLLPGNARCLPGVK  
QQLGTVRPLGMTSVFICLRGTEKEDLHLPSTNYVYYDTDMDQAMERYVSMPREEAAEHIPLLFFAFPSAKDPTW  
EDRFPGRSTMIMLIPTAYEWFEEWQAEKKGKRGSDYETFKNSFVEASMSVVLKLFPPQLEGKVESVTAGSPLTNQF  
YLAAPRGACYGADHDLGRLHPCVMASLRAQSPIPNLYLTGQDIFTCGLVGALQGALLCSSAILKRNLYSDLKNLD  
SRIRAQKKKN



94/615

**FIGURE 93**

GGGAAAGATGGCGGCGACTCTGGGACCCCTTGGGTCTGTGGCAGCAGTGGCGGCGATGTTTGTCTGGCTCGGGATGG  
GTCCAGGATGTTACTCCTTCTTCTTTTGTGGGTCTGGGCAGGGGCCACAGCAAGTCTGGGGCGGGTCAAACGTT  
CGAGTACTTGAAACGGGAGCACTCGCTGTCTGAAGCCCTACCAGGGTGTGGGCACAGGCAGTTCCCTCACTGTGGAA  
TCTGATGGGCAATGCCATGGTGATGACCCAGTATATCCGCCTTACCCAGATATGCAAAGTAAACAGGGTGCCTT  
GTGGAACCGGGTGCCATGTTTCCTGAGAGACTGGGAGTTGCAGGTGCACTTCAAATCCATGGACAAGGAAAGAA  
GAATCTGCATGGGGATGGCTTGGCAATCTGGTACACAAAGGATCGGATGCAGCCAGGGCCTGTGTTTGGAAACAT  
GGACAAATTTGTGGGGCTGGGAGTATTTGTAGACACCTACCCCAATGAGGAGAAGCAGCAAGAGCGGGTATTCCC  
CTACATCTCAGCCATGGTGAACAACGGCTCCCTCAGCTATGATCATGAGCGGATGGGCGGCCTACAGAGCTGGG  
AGGCTGCACAGCCATTGTCCGCAATCTTCATTACGACACCTTCTTGGTGATTCTCGCTACGTCAAGAGGCATTTGAC  
GATAATGATGGATATTGATGGCAAGCATGAGTGGAGGACTGCATTGAAGTGGCGGAGTCCGCCTGCCCCGCGG  
CTACTACTTCGGCACCTCCCTCCATCACTGGGGATCTCTCAGATAATCATGATGTCAATTTCTTGAAGTTGTTTGA  
ACTGACAGTGGAGAGAACCCCAAGAGGAAAAGCTCCATCGAGATGTGTTCTTGCCCTCAGTGGACAATATGAA  
GCTGCCTGAGATGACAGCTCCACTGCCGCCCTGAGTGGCCTGGCCCTCTTCTCATCGTCTTTTTCTCCCTGGT  
GTTTTCTGTATTTGCCATAGTCATTGGTATCACTCTACAACAAATGGCAGGAACAGAGCCGAAAGCGCTTCTA  
CTGAGCCCTCCTGCTGCCACCACTTTTGTGACTGTCAACCATGAGGTATGGAAGGAGCAGGCAGTGGCCTGAGCA  
TGCAGCCTGGAGAGTGTCTTGTCTCTAGCAGCTGGTTGGGGACTATATTTCTGCTACTGGAGTTTGAATGCAGG  
GACCCCGCATTTCCCATGGTTGTGCTGAGGGGACATCTAACTCTGGTCTGGGAAGCCACCCACCCAGGGCAATGCT  
GCTGTGATGTGCCTTTCCCTGCAGTCCTTCCATGTGGGAGCAGAGGTGTGAAGAGAATTTACGTGGTTGTGATGC  
CAAAATCACAGAACAGAATTTATAGCCCAGGCTGCCGTGTTGTTTGACTCAGAAGGCCCTTCTACTTCAGTTTT  
GAATCCACAAAGAATTAATACTGGTAACACCACAGGCTTTCTGACCATCCATTGTTGGGTTTTGCATTTGACC  
CAACCCTCTGCCTACCTGAGGAGCTTTCTTTGGAAACCAGGATGGAACTTCTTCCCTGCCTTACCTTCCTTTCA  
CTCCATTCAATTGTCCTCTCTGTGTGCAACCTGAGCTGGGAAAGGCATTTGGATGCCTCTCTGTTGGGGCCTGGGG  
CTGCAGAACACACCTGCGTTTCACTGGCCTTCATTAGGTGGCCCTAGGGAGATGGCTTTCTGCTTTGGATCACTG  
TTCCCTAGCATGGGTCTTGGGTCTATTGGCATGTCCATGGCCTTCCCAATCAAGTCTCTTCAGGCCCTCAGTGAA  
GTTTGGCTAAAGGTTGGTGTAAAAATCAAGAGAAGCCTGGAAGACATCATGGATGCCATGGATTAGCTGTGCAAC  
TGACCAGCTCCAGGTTTGATCAAACCAAAAGCAACATTTGTGATGTGGTCTGACCATGTGGAGATGTTTCTGGAC  
TTGCTAGAGCCTGCTTAGCTGCATGTTTTGTAGTTACGATTTTTTGAATCCCACTTTGAGTGCTGAAAGTGTAAG  
GAAGCTTTCTTCTTACACCTTGGGCTTGGATATTGCCCAGAGAAGAAATTTGGCTTTTTTTTTCTTAATGGACAA  
GAGACAGTTGCTGTTCTCATGTTCCAAGTCTGAGAGCAACAGACCCTCATCATCTGTGCCTGGAAGAGTTCAGTG  
TCATTGAGCAGCACAGCCTGAGTGTGGCCTCTGTCAACCCTTATTCCACTGCCTTATTTGACAAGGGGTTACAT  
GCTGCTCACCTTACTGCCCTGGGATTAAATCAGTTACAGGCCAGAGTCTCCTTGGAGGGCCTGGAAGTCTGAGTC  
CTCCTATGAACCTCTGTAGCCTAAATGAAATTTCTTAAATCACCGATGGAACCAAAAAAAAAAAAAAAAAAGGGCG  
CGCGGCACTCTAGAGTCGACCTGCAGTAGGGATAACAGGGTAATAAGCTTGGCCGCATGG

95/615

**FIGURE 94**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50911
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39711, pI: 8.70, NX(S/T): 1
MAATLGPIGSWQQWRRCLISARDGSRMLLLLLLLGSGQGPPQVVGAGQTFEYLKREHSLSKPYQGVGTGSSSLWNLM
GNAMVMTQYIIRLTPDMQSKQGALWNRVPCFLRDWELQVHFKIHGQKKNLHGDGLAIWYTKDRMQPGPVFGNMDK
FVGLGVFVDITYPNEEKQQERVFPYISAMVNNGSLSYDHERDGRPTLGGCTAIVRNLYDTFLVIRYVKRHLTIM
MDIDGKHEWRDCIEVPGVRLPRGYFYTSSITGDLSDNHDVISLKLFEITVERTPEEEKLHRDVFLPSVDNMKLP
EMTAPLPPLSGLALFLIVFFSLVFSVFAIVIGIILYNKWQEQRKREY
```

**Signal sequence:**  
amino acids 1-38

**Transmembrane domain:**  
amino acids 310-329

96/615

**FIGURE 95**

CCTGTGTTAAGCTGAGGTTTCCCCTAGATCTCGTATATCCCCAACACATACCTCCACGCACACACATCCCCAAGA  
ACCTCGAGCTCACACCAACAGACACACGCGCGCATACACACTCGCTCTCGCTTGTCCATCTCCCTCCCGGGGGAG  
CCGGCGCGCGCTCCCACCTTTGCCGCACACTCCGGCGAGCCGAGCCCGCAGCGCTCCAGGATTCTGCGGCTCGGA  
ACTCGGATTGCAGCTCTGAACCCCCATGGTGGTTTTTTTAAACACTTCTTTTCTCTCTCCTCGTTTGATTGC  
ACCGTTTCCATCTGGGGGCTAGAGGAGCAAGGCAGCAGCCTTCCCAGCCAGCCCTTGTGGCTTGCATCGTCCA  
TCTGGCTTATAAAAGTTTGTCTGAGCGCAGTCCAGAGGGCTGCGCTGCTCGTCCCCTCGGCTGGCAGAAGGGGGTG  
ACGCTGGGCAGCGGCGAGGAGCGCGCGCTGCTGCTGGCGGGCTTTCGGCTTGAGGGGCAAGGTGAAGAGCGCAC  
CGGCCGTGGGGTTTACCGAGCTGGATTGTATGTTGACCAATGCTTCTTGATCGGGGCTGTGATTCTTCCCCT  
CTTGGGGCTGCTGCTCTCCCTCCCGCGGGGCGGATGTGAAGGCTCGGAGCTCGGAGAGGTCCGCCAGGCGTA  
CGGTGCCAAGGGATTACGCCGCGGACATCCCCTACCAGGAGATCGCAGGGGAACACTTAAGAATCTGTCTCA  
GGAATATACATGCTGCACCACAGAAATGGAAGACAAGTTAAGCCAACAAAGCAAACCTCGAATTTGAAAACCTTGT  
GGAAGAGACAAGCCATTTTGTGCGCACCACCTTTTGTGTCCAGGCATAAGAAATTTGACGAATTTTCCGAGAGCT  
CCTGGAGAATGCAGAAAAGTCACTAAATGATATGTTTGTACGGACCTATGGCATGCTGTACATGCAGAATTCAGA  
AGTCTTCCAGGACCTCTTACAGAGCTGAAAAGGTACTACACTGGGGTAAATGTGAATCTGGAGGAAATGCTCAA  
TGACTTTTGGGCTCGGCTCCTGGAACGGATGTTTCAGCTGATAAACCCCTCAGTATCACTTCAGTGAAGACTACCT  
GGAATGTGTGAGCAAATACACTGACCAGCTCAAGCCATTTGGAGACGTGCCCCGAAACTGAAGATTACAGTTAC  
CCGCGCCTTCATTGCTGCCAGGACCTTTGTCCAGGGGCTGACTGTGGGCAGAGAAGTTGCAAACCGAGTTTCCAA  
GGTCAGCCCAACCCAGGGTGTATCCGTGCCCTCATGAAGATGCTGTACTGCCATACTGTGCGGGGCTTCCCAC  
TGTGAGGCCCTGCAACAACACTACTGTCTCAACGTATGAAGGGCTGCTTGGCAAATCAGGCTGACCTCGACACAGA  
GTGGAATCTGTTTATAGATGCAATGCTCTTGGTGCCAGAGCGACTGGAGGGGCCATTCAACATTGAGTCGGTCAT  
GGACCCGATAGATGTCAAGATTTCTGAAGCCATTATGAACATGCAAGAAAACAGCATGCAGGTGTCTGCAAAGGT  
CTTTCAGGGATGTGGTCAGCCCAAACCTGCTCCAGCCCTCAGATCTGCCGCTCAGCTCCTGAAAATTTTAATAC  
ACGTTTCAGGCCCTACAATCCTGAGGAAAGACCAACAACCTGCTGCAGGCACAAGCTTGGACCGGCTGGTCACAGA  
CATAAAGAGAAATTGAAGCTCTTAAAAAGGTCTGGTCAGCATTACCTACACTATCTGCAAGGACGAGAGCGT  
GACAGCGGGCACGTCCAACGAGGAGGAATGCTGGAACGGGCACAGCAAAGCCAGATACTTGCCTGAGATCATGAA  
TGATGGGCTCACCAACCAGATCAACAATCCCGAGGTGGATGTGGACATCACTCGGCCTGACACTTTCATCAGACA  
GCAGATTATGGCTCTCCGTGTGATGACCAACAACTAAAAACGCCCTACAATGGCAATGATGTCAATTTCCAGGA  
CACAGTGATGAATCCAGTGGCTCAGGGAGTGGCAGTGGGTGCATGGATGACGTGTGTCCCACGGAGTTTGAGTT  
TGTACACACAGAGGCCCCCGAGTGGATCCCGACCGGAGAGAGGTGGACTCTTCTGCAGCCAGCGTGGCCACTC  
CCTGCTCTCCTGGTCTCTCACCTGCATTGTCTTGGCACTGCAGAGACTGTGAGATAATCTTGGGTTTTTGGTCA  
GATGAAACTGCATTTTAGCTATCTGAATGGCCAACTCACTTCTTTTCTTACACTCTTGGACAATGGACCATGCCA  
CAAAACTTACCGTTTTCTATGAGAAGAGAGCAGTAATGCAATCTGCCTCCCTTTTTGTTTTCCCAAAGAGTACC  
GGGTGCCAGACTGAAGTCTTCTCTTTCTTTCAGCTATCTGTGGGACCTTGTATTCTAGAGAGAATTTCTTA  
CTCAAATTTTTCGTACCAGGAGATTTTCTTACCTTCATTGCTTTTATGCTGCAGAAGTAAAGGAATCTCACGTT  
GTGAGGGTTTTTTTTTCTCATTTAAAT

97/615

**FIGURE 96**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50914
><subunit 1 of 1, 555 aa, 1 stop
><MW: 62736, pI: 5.36, NX(S/T): 0
MPSWIGAVILPLLGLLLSLPAGADVKARSCGEVRQAYGAKGFSLADIPYQEIAGEHLRICPQEYTCCTTEMEDKL
SQQSKLEFENLVEETSHFVRTTFVSRHKKFDEFFRELLENAAEKSLNDMFVRTYGMLYMQNSEVFQDLFTELKRY
TGGNVNLEEMLNDFWARLLERMFQLINPQYHFSELYECVSKYTDQLKPFQDVPRKLKIQVTRAFIAARTFVQGL
TVGREVANRVSKVSPTPGCIRALMKMLYCPYCRGLPTVRPCNNYCLNVMKGCLANQADLDTEWNLFDAMLLVAE
RLEGPFNIESVMDPIDVKISEAIMNQENSMQVSAKVFQCGQPKPAPALRSARSAPENFNTRFRPYNPEERPTT
AAGTSLDRLVTDIKEKLKLSKKVWSALPYTICKDESVTAGTSNEEECWNGHASKARYLPEIMNDGLTNQINNPEVD
VDITRPDTFIRQQIMALRVMTNKLKNAYNGNDVNFQDTSDESSGSGSGSGCMDDVCPTFEFEFVTTEAPAVDPDRR
EVDSSAAQRGHSLLSWSLTCIVLALQRLCR
```

**Signal peptide:**  
amino acids 1-23

98/615

**FIGURE 97**

GGCGGCGTCCGTGAGGGGCTCCTTTGGGCAGGGGTTAGTGTGTTGGTGTCCCTGTCTTGCGTGATATTGACAAACTG  
AAGCTTTCCTGCACCACTGGACTTAAGGAAGAGTGTACTCGTAGGCGGACAGCTTTAGTGGCCGGCCGGCCGCTC  
TCATCCCCCGTAAGGAGCAGAGTCCTTTGTACTGACCAAGATGAGCAACATCTACATCCAGGAGCCTCCACGAA  
TGGGAAGGTTTTATTGAAACTACAGCTGGAGATATTGACATAGAGTTGTGGTCCAAAGAAGCTCCTAAAGCTTG  
CAGAAATTTTATCCAACCTTTGTTTGAAGCTTATTATGACAATACCATTTTTCATAGAGTTGTGCCTGGTTTCAT  
AGTCCAAGGCGGAGATCCTACTGGCACAGGGAGTGGTGGAGAGTCTATCTATGGAGCGCCATTCAAAGATGAATT  
TCATTACGGTTGCGTTTTAATCGGAGAGGACTGGTTGCCATGGCAAATGCTGGTTCTCATGATAATGGCAGCCA  
GTTTTTCTTCACACTGGGTCGAGCAGATGAACCTTAACAATAAGCATACCATCTTTGGAAAGGTTACAGGGGATAC  
AGTATATAACATGTTGCGACTGTCAGAAGTAGACATTGATGATGACGAAAGACCACATAATCCACACAAAATAAA  
AAGCTGTGAGGTTTTGTTAATCCTTTTGATGACATCATTCCAAGGGAAATTAAAAGGCTGAAAAAGAGAAACC  
AGAGGAGGAAGTAAAGAAATTGAAACCCAAAGGCACAAAAAATTTAGTTTACTTTTATTGGAGAGGAAGCTGA  
GGAAGAAGAGGAGGAAGTAAATCGAGTTAGTCAGAGCATGAAGGGCAAAAGCAAAAGTAGTCATGACTTGCTTAA  
GGATGATCCACATCTCAGTTCTGTTCCAGTTGTAGAAAGTGAAAAAGGTGATGCACCAGATTAGTTGATGATGG  
AGAAGATGAAAGTGCAGAGCATGATGAATATATTGATGGTGTGAAAAAGAACCTGATGAGAGAAAGAATTGCCAA  
AAAATTAAGGACACAAGTGCAGATGTTAAATCAGCTGGAGAAGGAGAAGTGGAGAAGAAATCAGTCAGCCG  
CAGTGAAGAGCTCAGAAAAAGCAAGACAATTAAAACGGGAACCTCTAGCAGCAAAACAAAAAAGTAGAAAA  
TGCAGCAAAACAAGCAGAAAAAGAAAGTGAAGAGGAAGAAGCCCTCCAGATGGTGTGTTGCCGAATACAGAAG  
AGAAAAGCAAAAGTATGAAGCTTTGAGGAAGCAACAGTCAAAGAAGGGAACCTCCCGGGAAGATCAGACCCCTTGC  
ACTGCTGAACAGTTTAAATCTAAACTCACTCAAGCAATTGCTGAAACACCTGAAATGACATTCTGAAACAGA  
AGTAGAAGATGATGAAGGATGGATGTCACATGTACTTCAGTTTGAGGATAAAAGCAGAAAAGTGAAGATGCAAG  
CATGCAAGACTCAGATACATTTGAAATCTATGATCCTCGGAATCCAGTGAATAAAAGAAGGAGGGAAGAAAGCAA  
AAAGCTGATGAGAGAGAAAAAGAAAGATATAATGAGAATAATGATAACCAGAACCTTGCTGGAAATGTGCCT  
ACAATGGCCTTGTAACAGCCATTGTTCCCAACAGCATCACTTAGGGGTGTGAAAAGAAGTATTTTTGAACCTGTT  
GTCTGGTTTTGAAAAACAATTATCTTTGTTTGCAAATTGTGGAATGATGTAAGCAAAATGCTTTGGTTACTGGTA  
CATGTGTTTTTTCCTAGCTGACCTTTTATATTGCTAAATCTGAAATAAATAAATTTCTTCCACAAAAA  
AAAAAAAAAAAAAAAAAAAA

99/615

**FIGURE 98**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50919
><subunit 1 of 1, 472 aa, 1 stop
><MW: 53847, pI: 5.75, NX(S/T): 2
MSNIYIQEPPTNGKVLLKTTAGDIDIELWSKEAPKACRNFIQLCLEAYYDNTIFHRVVPGFIVQGGDPTGTGSGG
ESYIGAPFKDEFHSRLRFNRRGLVAMANAGSHDNGSQFFFTLGRADELNKHTIFGKVTGDTVYNMLRLSEVDID
DDERPHNPHKIKSCEVLFPFDDIIPREIKRLKKEKPVEEVKKLKPCKGTKNFSLLSFGEEAEVEVNRSQSM
KGKSKSSHDLKDDPHLSSVPVVESEKGDAPDLVDDGEDESAEHDEYIDGDEKNLMRERIAKKLKKDTSANVKS
GEGEVEKKSVSRSEELRKEARQLKRELLAAKQKKVENAAKQAEKRSEEEAPPDGAVAEYRREKQKYEALRKQOS
KKGTSREDQTLALLNQFKSKLTQAIATPNDIPETEVDEDEGWMSHVLQFEDKSRKVKDASMQSDTFEIYDPR
NPVKNRRREESKMLREKKERR
```

**Important features:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 109-112 and 201-204

**Cyclophilin-type peptidyl-prolyl cis-trans isomerase signature.**

amino acids 49-66

**Homologous region to Cyclophilin-type peptidyl-prolyl cis-trans isomerase**

amino acids 96-140, 49-89 and 22-51

100/615

**FIGURE 99**

CTTTTCTGAGGAACCACAGCAATGAATGGCTTTGCATCCTTGCTTCGAAGAAACCAATTTATCCTCCTGGTACTA  
TTTCTTTTGCAAATTCAGAGTCTGGGTCTGGATATTGATAGCCGTCCTACCGCTGAAGTCTGTGCCACACACACA  
ATTTCAACCAGGACCCAAAGGAGATGATGGTGAAAAAGGAGATCCAGGAGAAGAGGGAAAGCATGGCAAAGTGCGA  
CGCATGGGGCCGAAAGGAATTAAAGGAGAACTGGGTGATATGGGAGATCAGGGCAATATTGGCAAGACTGGGCCC  
ATTGGGAAGAAGGGTGACAAAGGGGAAAAAGGTTTGCTTGGAATACCTGGAGAAAAAGGCAAAGCAGGTACTGTC  
TGTGATTGTGGAAGATACCGGAAATTTGTTGGACAACCTGGATATTAGTATTGCTCGGCTCAAGACATCTATGAAG  
TTTGTCAAGAATGTGATAGCAGGGATTAGGGAAACTGAAGAGAAATTTCTACTACATCGTGCAGGAAGAGAAGAAC  
TACAGGGAATCCCTAACCCACTGCAGGATTCGGGGTGGAATGCTAGCCATGCCCAAGGATGAAGCTGCCAACACA  
CTCATCGCTGACTATGTTGCCAAGAGTGGCTTCTTCGGGTGTTTCATTGGCGTGAATGACCTTGAAAGGGAGGGA  
CAGTACATGTCCACAGACAACACTCCACTGCAGAACTATAGCAACTGGAATGAGGGGGAACCCAGCGACCCCTAT  
GGTCATGAGGACTGTGTGGAGATGCTGAGCTCTGGCAGATGGAATGACACAGAGTGCCATCTTACCATGTACTTT  
GTCTGTGAGTTCATCAAGAAGAAAAAGTAACTTCCCTCATCCTACGTATTTGCTATTTTCTGTGACCGTCATTA  
CAGTTATTGTTATCCATCCTTTTTTTTCTGATTGTACTACATTTGATCTGAGTCAACATAGCTAGAAAATGCTAA  
ACTGAGGTATGGAGCCTCCATCATCAAAAAAAAAAAAAAAAAA

101/615

**FIGURE 100**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA50980
><subunit 1 of 1, 277 aa, 1 stop
><MW: 30645, pI: 7.47, NX(S/T): 2
MNGFASLLRRNQFILLVLFLLQIQSLGLDIDSRPTAEVCATHTISPGPKGDDGEKGDPEEGKHGKVGRMGPKGI
KGELGDMGDQGNIGKTGPIGKKGDKGEKGLLGIPGEKGKAGTVCDGGRYRKFBVGQLDISIARLKTSMKFVKNVIA
GIRETEEFYIIVQEEKNYRESLTHCRIRGGMLAMPKDEAANTLIADYVAKSGFFRVFIGVNDLEREGQYMSTDN
TPLQNYSNWNEGEPSDPYGHEDCVEMLSGRWNDTECHLTMYFVCEFIKKKK
```

**Signal peptide:**  
amino acids 1-25



102/615

**FIGURE 101**

GCAACCTCAGCTTCTAGTATCCAGACTCCAGCGCCGCCCCGGGCGCGGACCCCCAACCCCGACCCAGAGCTTCTCC  
AGCGGGCGGCGCAGCGAGCGAGGGCTCCCCGCCTTAACTTCCTCCGCGGGGGCCAGCCACCTTCGGGAGTCCGGGTT  
GCCACCTGCAAACCTCTCCGCCTTCTGCACCTGCCACCCCTGAGCCAGCGCGGGCCCCGAGCGAGT**CATG**GCCA  
ACGCGGGGCTGCAGCTGTTGGGCTTCATTCTCGCCTTCCTGGGATGGATCGGCGCCATCGTCAGCACTGCCCTGC  
CCCAGTGGAGGATTTACTCCTATGCCGGCGACAACATCGTGACCGCCCAGGCCATGTACGAGGGGCTGTGGATGT  
CCTGCGTGTCGCAGAGCACCGGGCAGATCCAGTGCAAAGTCTTTGACTCCTTGCTGAATCTGAGCAGCACATTGC  
AAGCAACCCGTGCCTTGATGGTGGTTGGCATCCTCCTGGGAGTGATAGCAATCTTTGTGGCCACCGTTGGCATGA  
AGTGTATGAAGTGCTTGGGAAGACGATGAGGTGCAGAAGATGAGGATGGCTGTCAATTGGGGGTGCGATATTTCTTC  
TTGCAGGTCTGGCTATTTTAGTTGCCACAGCATGGTATGGCAATAGAATCGTTCAGAATTCTATGACCCATGA  
CCCCAGTCAATGCCAGGTACGAATTTGGTCAGGCTCTCTTCACTGGCTGGGCTGCTGCTTCTCTCGCCTTCTGG  
GAGGTGCCCTACTTTGCTGTTCTGTCCCCGAAAAACAACCTCTTACCCAACACCAAGGCCCTATCCAAAACCTG  
CACCTTCCAGCGGGAAAGACTACGTG**TGAC**CACAGAGGCAAAAGGAGAAAATCATGTTGAAACAAACCGAAAAATGG  
ACATTGAGATACTATCATTAAACATTAGGACCTTAGAATTTTGGGTATTGTAATCTGAAGTATGGTATTACAAAAC  
AAACAAACAAACAAAAAACCATGTGTTAAATACTCAGTGCTAAACATGGCTTAATCTTATTTTATCTTCTTTTC  
CTCAATATAGGAGGAAGATTTTCCATTTGTATTACTGCTTCCCATTGAGTAATCATACTCAAATGGGGGAAGG  
GGTGCTCCTTAAATATATATAGATATGTATATATACATGTTTTTCTATTAAAAATAGACAGTAAATACTATTCT  
CATTATGTTGATACTAGCATACTTAAATATCTCTAAATAGGTAAATGTATTTAATTCCATATTGATGAAGATG  
TTTATTGGTATATTTTCTTTTTCGTCCTTATATACATATGTAACAGTCAAATATCATTTACTCTTCTTCATTAGC  
TTTGGGTGCCTTTGCCACAAGACCTAGCCTAATTTACCAAGGATGAATCTTTCAATTCTTCATGCGTGCCCTTT  
TCATATACTTATTTTATTTTACCATAATCTTATAGCACTTGCATCGTTATTAAGCCCTTATTTGTTTTGTGTT  
TCATTGGTCTCTATCTCCTGAATCTAACACATTTCATAGCCTACATTTTAGTTTCTAAAGCCAAGAAGAAATTTAT  
TACAAATCAGAACTTTGGAGGCAAATCTTTCTGCATGACCAAAGTGATAAATTCCTGTTGACCTTCCACACAAT  
CCCTGTACTCTGACCCATAGCACTCTTGTGTTGCTTTGAAAATATTTGTCCAATTGAGTAGCTGCATGCTGTTCCC  
CCAGGTGTTGTAACACAACCTTTATTGATTGAATTTTTAAGCTACTTATTCATAGTTTTATATCCCCCTAAACTAC  
CTTTTTGTTCCCATTCCTTAATTGTATTGTTTTCCCAAGTGTAATTATCATGCGTTTTATATCTTCTCTAATAAG  
GTGTGGTCTGTTTGTCTGAACAAAGTGCTAGACTTTCTGGAGTGATAATCTGGTGACAAATATCTCTCTGTAGC  
TGTAAGCAAGTCACTTAATCTTTCTACCTCTTTTTCTATCTGCCAAATTGAGATAATGATACTTAACCAAGTTAG  
AAGAGGTAGTGTAATATTAATTAGTTTATATTACTCTTATTCTTTGAACATGAACATGCCTATGTAGTGTCTT  
TATTTGCTCAGCTGGCTGAGACACTGAAGAAGTCACTGAACAAAACCTACACACGTACCTTCATGTGATTCACTG  
CCTTCCTCTCTCTACCAGTCTATTTCCACTGAACAAAACCTACACACATACCTTCATGTGGTTTCAGTGCCTTCCT  
CTCTCTACCAGTCTATTTCCACTGAACAAAACCTACGCACATACCTTCATGTGGCTCAGTGCCTTCCTCTCTCTA  
CCAGTCTATTTCCATTCTTTAGCTGTGTCTGACATGTTTGTGCTCTGTTCCATTTTAAACACTGCTCTTACTTT  
TCCAGTCTGTACAGAATGCTATTTCACTTGAGCAAGATGATGTAATGGAAAGGGTGTGGCACTGGTGTCTGGAG  
ACCTGGATTTGAGTCTTGGTGCTATCAATCACCGTCTGTGTTTGAGCAAGGCATTTGGCTGCTGTAAGCTTATTG  
CTTCATCTGTAAGCGGTGGTTTGTAAATTCCTGATCTTCCACCTCACAGTGATGTTGTGGGGATCCAGTGAGATA  
GAATACATGTAAGTGTTGTTTGTAAATTTAAAAAGTGCTATACTAAGGGAAAGAATTGAGGAATTAACGCATAC  
GTTTTGGTGTGCTTTTCAAATGTTTAAAAATAAAAAAATGTTAAG

103/615

**FIGURE 102**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52185
><subunit 1 of 1, 211 aa, 1 stop
><MW: 22744, pI: 8.51, NX(S/T): 1
MANAGLQLLGFI LAF LGWIGAIVSTALPQWRIYSYAGDNIVTAQAMYEG LWMSCVSQSTGQIQCKVFDSLNLSS
TLQATRALMVVGILLGVIAIFVATVGMKCMKCLEDDEVQKMRMAVIGGAIFLLAGLAILVATAWYGNRIVQEFYD
PMTFVNARYEFGQALFTGWAAASLCLLGALLCCSCPRKTTSYPTPRPYPKPAPSSGKD YV
```

**Important features:****Signal peptide:**

amino acids 1-21

**Transmembrane domains:**

amino acids 82-102, 118-142 and 161-187

**N-glycosylation site.**

amino acids 72-75

**PMP-22 / EMP / MP20 family proteins**

amino acids 70-111

**ABC-2 type transport system integral membrane protein**

amino acids 119-133

104/615

**FIGURE 103**

CCCACGCGTCCGCGGACGCGTGGGCTGGACCCAGGTCTGGAGCGAATTCCAGCCTGCAGGGCTGATAAGCGAGG  
CATTAGTGAGATTGAGAGAGACTTTACCCCGCGTGGTGGTTGGAGGGCGCGCAGTAGAGCAGCAGCACAGGCGC  
GGGTCCCGGGAGGCGCGCTCTGCTCGCGCCGAGATGTGGAATCTCCTTCACGAAACCGACTCGGCTGTGGCCACC  
GCGCGCCGCGCGCTGGCTGTGCGCTGGGGCGTGGTGTGGCGGGTGGCTTCTTCTCCTCGGCTTCCTCTTC  
GGGTGGTTTATAAAATCCTCCAATGAAGCTACTAACATTACTCCAAAGCATAATATGAAAGCATTTTTGGATGAA  
TTGAAAGCTGAGAACATCAAGAAGTTCTTACATAATTTTACACAGATACCACATTTAGCAGGAACAGAACAAAAC  
TTTCAGCTTGCAAAGCAAATTCATCCCAGTGGAAGAATTTGGCCTGGATTCTGTTGAGCTAGCTCATTATGAT  
GTCCTGTTGTCCTACCCAAATAAGACTCATCCCACTACATCTCAATAATTAATGAAGATGGAATGAGATTTTC  
AACACATCATTTATTTGAACCACCTCCTCCAGGATATGAAAATGTTTCGGATATTGTACCACCTTTCAGTGCTTTC  
TCTCCTCAAGGAATGCCAGAGGGCGATCTAGTGTATGTTAACTATGCACGAACTGAAGACTTCTTTAAATTGGAA  
CGGGACATGAAAATCAATTGCTCTGGGAAAATTGTAATTGCCAGATATGGGAAAGTTTTCAGAGGAAATAAGGTT  
AAAAATGCCAGCTGGCAGGGGCCAAAGGAGTCATTCTCTACTCCGACCCTGCTGACTACTTTGCTCCTGGGGTG  
AAGTCCATATCCAGACGGTTGGAATCTTCTGGAGGTGGTGTCCAGCGTGGAAATATCCTAAATCTGAATGGTGCA  
GGAGACCCTCTCACACCAGGTTACCCAGCAAATGAATATGCTTATAGGCGTGGAAATGCAGAGGCTGTTGGTCTT  
CCAAGTATTCTGTTTCATCCAATTGGATACCTATGATGCACAGAAGCTCCTAGAAAAAATGGGTGGCTCAGCACC  
CCAGATAGCAGCTGGAGAGGAAGTCTCAAAGTGCCCTACAATGTTGGACCTGGCTTTTACTGGAAAATTTTCTACA  
CAAAAAGTCAAGATGCACATCCACTCTACCAATGAAGTGACGAGAATTTACAATGTGATAGGTACTCTCAGAGGA  
GCAGTGGAACCAGACAGATATGTCATTCTGGGAGGTCACCGGGACTCATGGGTGTTTGGTGGTATTGACCCTCAG  
AGTGGAGCAGCTGTTGTTTCATGAAATTGTGAGGAGCTTTGGAACACTGAAAAGGAAGGGTGGAGACCTAGAAGA  
ACAATTTTGTGTTGCAAGCTGGGATGCAGAAGAATTTGGTCTTCTTGGTTCTACTGAGTGGGCAGAGGAGAATTC  
AGACTCCTTCAAGAGCGTGGCGTGGCTTATATTAATGCTGACTCATCTATAGAAGGAACTACACTCTGAGAGTT  
GATTGTACACCGCTGATGTACAGCTTGGTACACAACCTAACAAAAGAGCTGAAAAGCCCTGATGAAGGCTTTGAA  
GGCAAATCTCTTTATGAAAGTTGGACTAAAAAAGTCTTCCCCAGAGTTCAGTGGCATGCCAGGATAAGCAAA  
TTGGGATCTGGAAATGATTTTGGAGGTGTTCTTCCAACGACTTGGAATTGCTTCAGGCAGAGCACGGTATACTAAA  
AATTGGGAAACAAACAAATTCAGCGGCTATCCACTGTATCACAGTGTCTATGAAACATATGAGTTGGTGGAAAAG  
TTTTATGATCCAATGTTTAAATATCACCTCACTGTGGCCCAGGTTTCAGAGAGGGATGGTGGTTTGGCTAGCCAAT  
TCCATAGTGCTCCCTTTTGGATTGTGCGAGATTATGCTGTAGTTTTTAAGAAAGTATGCTGACAAAATCTACAGTATT  
TCTATGAAACATCCACAGGAAATGAAGACATACAGTGTATCATTTGATTCACTTTTTTCTGCAGTAAAGAATTTT  
ACAGAAATTGCTTCCAAGTTTCAGTGAGAGACTCCAGGACTTTGACAAAAGCAACCCAATAGTATTAAGAATGATG  
AATGATCAACTCATGTTTCTGGAAAGAGCATTATTTGATCCATTAGGGTTACCAGACAGGCCTTTTTATAGGCAT  
GTCATCTATGCTCCAAGCAGCCACAACAAGTATGCAGGGGAGTCATTCCCAGGAATTTATGATGCTCTGTTTGAT  
ATTGAAAGCAAAGTGGACCCCTCCAAGGCCTGGGGAGAAGTGAAGAGACAGATTTATGTTGCAGCCTTCACAGTG  
CAGGCAGCTGCAGAGACTTTGAGTGAAGTAGCCTAAGAGGATTTTTTAGAGAATCCGTATTGAATTTGTGTGGTA  
TGTCACTCAGAAAGAATCGTAATGGGTATATTGATAAATTTTAAATTTGGTATATTTGAAATAAAGTTGAATATT  
ATATATAA

105/615

**FIGURE 104**

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA52756
><subunit 1 of 1, 750 aa, 1 stop
><MW: 84305, pI: 6.93, NX(S/T): 10
MWNLHETDSAVATARRPRWLCAGALVLAGGFFLLGFLFGWFIKSSNEATNITPKHNMKAFLDELKAENIKKFLH
NFTQIPHLAGTEQNFQLAKQIQSQWKEFGLDSVELAHYDVLLSYPNKTHPNYISIINEDGNEIFNTSLFEPPPPPG
YENVSDIVPPFSAFSPQGMPEGDLVYVNYARTEDFFKLERDMKINCSGKIVIARYGKVFRGNKVNAQLAGAKGV
ILYSDPADYFAPGVKSYPDGWNLPGGGVQQRGNILNLNGAGDPLTPGYPANAYARRGIAEAVGLPSIPVHPIGYY
DAQKLEKMGGSAPPDSSWRGSLKVPYNVGPFTGNFSTQKVKMHIHSTNEVTRIYNVIGTLRGAVEPDRYVILG
GHRDSWVFGGIDPQSGAAVVHEIVRSFGTLKKEGWRPRRTILFASWDAEEFGLLGSTEWAEENSRLQERGVAYI
NADSSIEGNYTLRVDCPLMYSLVHNLTKELKSPDEGFEGKSLYESWTKKSPSPEFSGMPRISKLGSGNDFEVFF
QRLGIASGRARYTKNWETNKFSGYPLYHSVYETYELVEKFYDPMFKYHLTVAQVRGGMVFELANSIVLPFDCRDY
AVVLRKYADKIYSISMKHPQEMKTYSVSFDLSFAVKNFTEIASKFSERLQDFDKSNPIVLRMMNDQLMFLERAF
IDPLGLPDRPFYRHVIYAPSSHNKYAGESFPGIYDALFDIESKVDPKAWGEVKRQIYVAAFTVQAAAETLSEVA
```

**Signal sequence:**

amino acids 1-40

**N-glycosylation sites.**amino acids 76-80, 121-125, 140-144, 153-157, 195-199, 336-340, 459-463,  
476-480, 638-642**Tyrosine kinase phosphorylation sites.**

amino acids 363-372, 605-613, 606-613, 617-626

**N-myristoylation sites.**amino acids 85-91, 168-174, 252-258, 256-262, 282-288, 335-341, 360-366,  
427-433, 529-535, 707-713

106/615

**FIGURE 105**

TGAAGAGTAATAGTTGGAATCAAAAGAGTCAACGCAATGAAGTGTATTTACTGCTGCGTTTTATGTTGGGAATT  
CCTCTCCTATGGCCTTGTCTTGGAGCAACAGAAACTCTCAAACAAAGAAAGTCAAGCAGCCAGTGGCATCTCAT  
TTGAGAGTGAAGCGTGGCTGGGTGTGGAACCAATTTTTTGTACCAGAGGAAATGAATACGACTAGTCATCACATC  
GGCCAGCTAAGATCTGATTTAGACAATGGAAACAATTTCTTCCAGTACAAGCTTTTGGGAGCTGGAGCTGGAAGT  
ACTTTTATCATTGATGAAAGAACAGGTGACATATATGCCATACAGAAGCTTGATAGAGAGGAGCGATCCCTCTAC  
ATCTTAAGAGCCCAGGTAATAGACATCGTACTGGAAGGGCTGTGGAACCTGAGTCTGAGTTTGTTCATCAAAGTT  
TCGGATATCAATGACAATGAACCAAAATTCCTAGATGAACCTTATGAGGCCATTGTACCAGAGATGTCTCCAGAA  
GGAACATTAGTTATCCAGGTGACAGCAAGTGATGCTGACGATCCCTCAAGTGGTAATAATGCTCGTCTCCTCTAC  
AGCTTACTTCAAGGCCAGCCATATTTTTCTGTTGAACCAACAACAGGAGTCATAAGAATATCTTCTAAAATGGAT  
AGAGAACTGCAAGATGAGTATTGGGTAATCATTCAAGCCAAGGACATGATTGGTCAGCCAGGAGCGTTGTCTGGA  
ACAACAAGTGTATTAATAAATTTTCCAGATGTTAATGACAATAAGCCTATATTTAAAGAAAAGTTTATACCGCTTG  
ACTGTCTCTGAATCTGACCCACTGGGACTTCTATAGGAACAATCATGGCATATGATAATGACATAGGAGAGAAT  
GCAGAAATGGATTACAGCATTGAAGAGGATGATTCGCAAAACATTTGACATTATTACTAATCATGAACTCAAGAA  
GGAATAGTTATATTAATAAAGAAAGTGGATTTTGAGCACCAGAACCACTACGGTATTAGAGCAAAAGTTAAAAAC  
CATCATGTTTCCCTGAGCAGCTCATGAAGTACCACACTGAGGCTTCCACCCTTTCATTAAGATCCAGGTGGAAGAT  
GTTGATGAGCCTCCTCTTTTCTCCTTCCATATTATGTATTTGAAGTTTTTTGAAGAAACCCACAGGGATCATT  
GTAGGCGTGGTGTCTGCCACAGACCCAGACAATAGGAAATCTCCTATCAGGTATTCTATTACTAGGAGCAAAGTG  
TTCAATATCAATGATAATGGTACAATCACTACAAGTAAGTCACTGGATCGTGAAATCAGTGCTTGGTACAACCTA  
AGTATTACAGCCACAGAAAAATACAATATAGAACAGATCTCTTCGATCCCACTGTATGTGCAAGTTCTTAACATC  
AATGATCATGCTCCTGAGTTCTCTCAATACTATGAGACTTATGTTTGTGAAATGCAGGCTCTGGTCAGGTAATT  
CAGACTATCAGTGCAGTGGATAGAGATGAATCCATAGAAGAGCACCATTTTACTTTAATCTATCTGTAGAAGAC  
ACTAACAATTCAAGTTTTTACAATCATAGATAATCAAGATAACACAGCTGTCATTTTGACTAATAGAAGTGGTTTT  
AACCTTCAAGAAGAACCTGTCTTCTACATCTCCATCTTAATTGCCGACAATGGAATCCCGTCACTTACAAGTACA  
AACACCCCTTACCATCCATGTCTGTGACTGTGGTGACAGTGGGAGCACACAGACCTGCCAGTACCAGGAGCTTGTG  
CTTTCCATGGGATTCAAGACAGAAGTTATCATTGCTATTCTCATTTGCATTATGATCATATTTGGGTTTTATTTTT  
TTGACTTTGGGTTTAAAAACAACGGAGAAAACAGATTCTATTTCTGAGAAAAGTGAAGATTTTCAAGAGAAATATA  
TTCCAATATGATGATGAAGGGGTGGAGAAGAAGATACAGAGGCCCTTTGATATAGCAGAGCTGAGGAGTAGTACC  
ATAATGCGGGAACGCAAGACTCGGAAAACCACAAGCGCTGAGATCAGGAGCCTATACAGGCAGTCTTTGCAAGTT  
GGCCCCGACAGTGCCATATTCAGGAAATTCATCTGGAAGGCTCGAAGAAGCTAATACTGATCCGTGTGCCCT  
CCTTTTGAATCCCTCCAGACCTACGCTTTTGAGGGAACAGGGTCATTAGCTGGATCCCTGAGCTCCTTAGAATCA  
GCAGTCTCTGATCAGGATGAAAGCTATGATTACCTTAATGAGTTGGGACCTCGCTTTAAAGATTAGCATGCATG  
TTTGGTTCTGCAGTGCAGTCAAATAATTAGGGCTTTTTACCATCAAATTTTTAAAGTGCTAATGTGTATTTCGA  
ACCCAATGGTAGTCTTAAAGAGTTTTGTGCCCTGGCTCTATGGCGGGGAAAGCCCTAGTCTATGGAGTTTTCTGA  
TTTTCCCTGGAGTAAATACTCCATGGTTATTTTAAAGCTACCTACATGCTGTCATTGAACAGAGATGTGGGAGAAA  
TGTAACAATCAGCTCACAGGCATCAATACAACCAGATTTGAAGTAAATAATGTAGGAAGATATTTAAAGTAGA  
TGAGAGGACACAAGATGTAGTCGATCCTTATGCGATTATATCATTATTTACTTAGGAAAGAGTAAAAATACCAAA  
CGAGAAAATTTAAAGGAGCAAAAATTTGCAAGTCAAATAGAAATGTACAAATCGAGATAACATTTACATTTCTAT  
CATATTGACATGAAAATTGAAAATGTATAGTCAGAGAAATTTTCATGAATTATTCCATGAAGTATTTCTCCTTT  
ATTTAAA

107/615

**FIGURE 106**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53906
><subunit 1 of 1, 772 aa, 1 stop
><MW: 87002, pI: 4.64, NX(S/T): 8
MNCYLLLRFMLGIPLLPCLGATENSQTKKVKQPVRSRLRVKRGWVWNQFFVPEEMNTTSHHIGQLRSDLNNGNN
SFQYKLLGAGAGSTFIIDERTGDIYAIQKLDREERSLYILRAQVIDIATGRAVEPESEFVIKVSINDNEPKFLD
EPYEAIVPEMSPEGTLVIQVTASDADDPSSGNARLLYSLLQGQPYFSVEPTTGVRIRISSKMDRELQDEYWVITQ
AKDMIGQPGALSGTTSVLIKLSDVNDNKPIFKESLYRLTVSESAPTGTSGITIMAYDNDIGENAEMDYSIEEDDS
QTFDIITNHETQEGIVILKKKVD FEHQNHYGIRAKVKNHHVPEQLMKYHTEASTTFIKIQVEDVDEPPLFLLPY
VFEVFEETPQGSFVGVVSATDPDNRKSPIRYSITRSKVFENINDNGTITTSNSLDREISAWYNLSITATEKYNIEQ
ISSIPLYVQVLNINDHAPEFSQYYETYVCENAGSGQVIQTI SAVDRDESIEHHFYFNLSVEDTNNSSFTIIDNQ
DNTAVILTNRGTGFNLQEEPVFYISILIA DNGIPSLTSTNTLTIHVCD CGDSGSTQTCQYQELVLSMGFKTEVIA
ILICIMIIFGFIFLTGLKQRRKQILFPEKSEDFRENIFQYDDEGGGEEDTEAFDIAELRSSTIMRERKTRKTTS
AEIRSLYRQSLQVGPDSAIFRKFILEKLEEANTDPCAPPFDSLQTYAFECTGSLAGSLSSLESAVSDQDESYDYL
NELGPRFKRLACMFGSAVQSNN
```

**Important features:****Signal peptide:**

amino acids 1-21

**Transmembrane domain:**

amino acids 597-617

**N-glycosylation sites.**

amino acids 57-60, 74-77, 419-423, 437-440, 508-511, 515-518, 516-519 and 534-537

**Cadherins extracellular repeated domain signature.**

amino acids 136-146 and 244-254

108/615

**FIGURE 107**

ATCTGGTTGAACTACTTAAGCTTAATTTGTTAAACTCCGGTAAGTACCTAGCCCACATGATTTGACTCAGAGATT  
CTCTTTTGTCCACAGACAGTCATCTCAGGGGCAGAAAGAAAGAGCTCCCAAATGCTATATCTATTTCAGGGGGCTC  
TCAAGAACAATGGAATATCATCTGATTTAGAAAATTTGGATGAAGATGGATATACTCAATTACACTTCGACTCT  
CAAAGCAATACCAGGATAGCTGTTGTTTCAGAGAAAGGATCGTGTGCTGCATCTCCTCCTTGGCGCCTCATTGCT  
GTAATTTTGGGAATCCTATGCTTGGTAATACTGGTGATAGCTGTGGTCCTGGGTACCATGGGGGTTCTTTCCAGC  
CCTTGTCCTCCTAATTGGATTATATATGAGAAGAGCTGTTATCTATTTCAGCATGTCACTAAATTCCTGGGATGGA  
AGTAAAAGACAATGCTGGCAACTGGGCTCTAATCTCCTAAAGATAGACAGCTCAAATGAATTGGGATTTATAGTA  
AAACAAGTGTCTTCCCAACCTGATAATTCATTTTGGATAGGCCTTTCTCGGGCCCCAGACTGAGGTACCATGGCTC  
TGGGAGGATGGATCAACATTCTCTTAACCTTATTTTCAGATCAGAACCACAGCTACCCAAGAAAACCCATCTCCA  
AATTGTGTATGGATTACAGTGTCACTCATTATGACCAACTGTGTAGTGTGCCCTCATATAGTATTTGTGAGAAG  
AAGTTTTCAATGTAAGAGGAAGGGTGGAGAAGGAGAGAGAAATATGTGAGGTAGTAAGGAGGACAGAAAACAGAA  
CAGAAAAGAGTAACAGCTGAGGTCAAGATAAATGCAGAAAATGTTTAGAGAGCTTGGCCAACCTGTAATCTTAACC  
AAGAAATTGAAGGGAGAGGCTGTGATTTCTGTATTTGTGACCTACAGGTAGGCTAGTATTATTTTTCTAGTTAG  
TAGATCCCTAGACATGGAATCAGGGCAGCCAAGCTTGAGTTTTTATTTTTTATTTATTTATTTTTTGGATAGG  
GTCTCACTTTGTTACCCAGGCTGGAGTGCAGTGGCACAATCTCGACTCACTGCAGCTATCTCTCGCCTCAGCCCC  
TCAAGTAGCTGGGACTACAGGTGCATGCCACCATGCCAGGCTAATTTTTTGGTGTTTTTTGTAGAGACTGGGTTTT  
GCCATGTTGACCAAGCTGGTCTCTAACTCCTGGGCTTAAGTGATCTGCCCCGCTTGGCCTCCCAAAGTGCTGGGA  
TTACAGATGTGAGCCACCACACCTGGCCCCAAGCTTGAATTTTCATTCTGCCATTGACTTGGCATTACCTTGGG  
TAAGCCATAAGCGAATCTTAATTTCTGGCTCTATCAGAGTTGTTTCATGCTCAACAATGCCATTGAAGTGCACGG  
TGTGTTGCCACGATTTGACCCTCAACTTCTAGCAGTATATCAGTTATGAAGTGAAGGTGAAATATATTTCTGAAT  
AGCTAAATGAAGAAATGGGAAAAAATCTTCACCACAGTCAGAGCAATTTTATTATTTTCATCAGTATGATCATAA  
TTATGATTATCATCTTAGTAAAAAGCAGGAACCTCCTACTTTTTCTTTATCAATTAAATAGCTCAGAGAGTACATC  
TGCCATATCTCTAATAGAATCTTTTTTTTTTTTTTTTTTTTTTTTGGACAGAGTTTCGCTCTTGTTGCCAGGCTG  
GAGTGCAACGGCAGATCTCGGCTCACCGCAACCTCCGCCCCCTGGGTTCAAGCAATTCCTGCTCAGCCTCC  
CAAGTAGCTGGGATTACAGTCAGGCACCACCACCCCGCTAATTTTGTATTTTTTTAGTAGAGACAGGGTTTCT  
CCATGTGGTCAGGCTAGTCCCGAACTCCTGACCTCAAGTGATCTGCCTGCCTCGGCCTCCCAAGTGCTGGGATT  
ACAGGCGTGAGCCACTGCACCCAGCCTAGAATCTTGATAATATGTAATTGTAGGGAACTGCTCTCATAGGAAA  
GTTTTCTGCTTTTTTAAATACAAAATACATAAAAATACATAAAATCTGATGATGAATATAAAAAAGTAACCAACC  
TCATTGGAACAAGTATTAACATTTTGAATATGTTTTATTAGTTTTGTGATGTACTGTTTTACAATTTTACCAT  
TTTTTTCAGTAATTACTGTAAAATGGTATTATTGGAATGAACTATATTTCTCATGTGCTGATTTGTCTTATTT  
TTTTCATACTTTCCCACTGGTGCTATTTTTATTTCCAATGGATATTTCTGTATTACTAGGGAGGCATTTACAGTC  
CTCTAATGTTGATTAATATGTGAAAAGAAATGTACCAATTTTACTAAATTATGCAGTTTAAATGGATGATTTT  
ATGTTATGTGGATTTCAATTAATAAAAAAACTCTTATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

109/615

**FIGURE 108**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53912  
<subunit 1 of 1, 201 aa, 1 stop  
<MW: 22563, pI: 4.87, NX(S/T): 1  
MEYHPDLENLDEEDGYTQLHFDSQSNTRIAVVSEKGSAAASPPWRLLIAVILGILCLVILVIAVVLGTMGVLSSPCP  
PNWIIYEKSCYLFMSLSNSWDGSKRQCWQLGSNLLKIDSSNELGFIVKQVSSQPDNSFWIGLSRPQTEVPWLWED  
GSTFSSNLFQIRTTATQENPSPNCVWIHVSVIYDQLCSVPSYSICEKKFSM

**Important features:****Type II transmembrane domain:**

amino acids 45-65

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 197-200

**N-myristoylation sites.**

amino acids 35-40 and 151-156

**Homologous region to LDL receptor**

amino acids 34-67 and 70-200.



110/615

**FIGURE 109**

CTGCAAGTTGTTAACGCCTAACACACAAGTATGTTAGGCTTCCACCAAAGTCCTCAATATACCTGAATACGCACA  
ATATCTTAACTCTTCATATTTGGTTTTGGGATCTGCTTTGAGGTCCCATCTTCATTTAAAAAAAATACAGAGAC  
CTACCTACCCGTACGCATACATATGTGTATATATATGTAAACTAGACAAAGATCGCAGATCATAAAGCAAG  
CTCTGCTTTTAGTTTTCCAAGAAGATTACAAAGAATTTAGAGATGTTATTTGTCAAGATCCCTGTCCGATTTCATGCCCT  
TTGGGTTACGGTGTCTCAGTGATGCAGCCCTACCCTTTGGTTTGGGGACATTATGATTTGTGTAAGACTCAGAT  
TTACACGGAAGAAGGGAAAGTTTGGGATTACATGGCCTGCCAGCCGAATCCACGGACATGACAAAATATCTGAA  
AGTGAAACTCGATCCTCCGGATATTACCTGTGGAGACCCTCCTGAGACGTTCTGTGCAATGGGCAATCCCTACAT  
GTGCAATAATGAGTGTGATGCGAGTACCCCTGAGCTGGCACACCCCTGAGCTGATGTTTGATTTTGAAGGAAG  
ACATCCCTCCACATTTTGGCAGTCTGCCACTTGAAGGAGTATCCCAAGCCTCTCCAGGTAAACATCACTCTGTC  
TTGGAGCAAACCATTGAGCTAACAGACAACATAGTTATTACCTTTGAATCTGGGCGTCCAGACCAAATGATCCT  
GGAGAAGTCTCTCGATTATGGACGAACATGGCAGCCCTATCAGTATTATGCCACAGACTGCTTAGATGCTTTTCA  
CATGGATCCTAAATCCGTGAAGGATTTATCACAGCATAACGGTCTTAGAAATCATTTGCACAGAAGAGTACTCAAC  
AGGGTATACAACAAATAGCAAAATAATCCACTTTGAAATCAAAGACAGGTTGCGGCTTTTTGCTGGACCTCGCCT  
ACGCAATATGGCTTCCCTCTACGGACAGCTGGATACAACCAAGAACTCAGAGATTTCTTTACAGTCACAGACCT  
GAGGATAAGGCTGTTAAGACCAGCCGTGGGGAAATATTTGTAGATGAGCTACACTTGGCAGCTACTTTTACGC  
GATCTCAGACATAAAGGTGCGAGGAAGGTGCAAGTGTAACTCTCCATGCCACTGTATGTGTGTATGACAACAGCAA  
ATTGACATGCCAATGTGAGCACAACTACAGGTCCAGACTGTGGGAAATGCAAGAAGAATTATCAGGGCCGACC  
TTGGAGTCCAGGCTCCTATCTCCCATCCCCAAGGCACGTGCAAATACCTGTATCCCCAGTATTTCCAGTATTGG  
TACGAATGTCTGCGACAACGAGCTCCTGCACTGCCAGAACGGAGGACGTGCCACAACAACGTGCGCTGCCTGTG  
CCCCGCCGCATACACGGGCATCCTCTGCGAGAAGCTGCGGTGCGAGGAGGCTGGCAGCTGCGGCTCCGACTCTGG  
CCAGGGCGCGCCCCCGCACGGCACCCAGCGCTGCTGCTGCTGACCACGCTGCTGGGAACCGCCAGCCCCCTGGT  
GTTCTAGGTGTACCTCCAGCCACACCGGACGGCCCTGTGCCGTGGGGAAGCAGACACAACCCAAACATTTGCTA  
CTAACATAGGAAACACACACATACAGACACCCCCACTCAGACAGTGTACAAACTAAGAAGGCCCTAACTGAACTAA  
GCCATATTTATCACCCGTGGACAGCACATCCGAGTCAAGACTGTTAATTTCTGACTCCAGAGGAGTTGGCAGCTG  
TTGATATTATCACTGCAAATCACATTGCCAGCTGCAGAGCATATTGTGGATTGAAAGGCTGCGACAGCCCCCA  
AACAGGAAAGACAAAAACAAACAAATCAACCGACCTAAAAACATTGGCTACTCTAGCGTGGTGCGCCCTAGTAC  
GACTCCGCCCAGTGTGTGGACCAACCAATAGCATTCTTTGCTGTGAGGTGCATTGTGGGCATAAGGAAATCTGT  
TACAAGCTGCCATATTGGCCTGCTTCCGTCCCTGAATCCCTTCCAACCTGTGCTTTAGTGAACGTTGCTCTGTAA  
CCCTCGTTGGTTGAAAGATTTCTTTGTCTGATGTTAGTGATGCACATGTGTAACAGCCCCCTTAAAAAGCGCAAG  
CCAGTCATACCCCTGTATATCTTAGCAGCACTGAGTCCAGTGCGAGCACACCCCACTATACAAGAGTGGCTATA  
GGAAAAAGAAAGTGTATCTATCCTTTTGTATTCAAATGAAGTTATTTTCTTGAACACTGTAAATATGTAGATT  
TTTTGTATTATTGCCAATTTGTGTTACCAGACAATCTGTTAATGTATCTAATTCGAATCAGCAAAGACTGACATT  
TTATTTTGTCTCTTTCTGTTCTGTTTGTCTTCACTGTGCAGAGATTTCTCTGTAAGGGCAACGAACGTGCTGGCA  
TCAAAGAATATCAGTTTACATATATAACAAGTGTAAAGATTCCACCAAAGGACATTCTAAATGTTTTCTTGT  
GCTTTAACACTGGAAGATTTAAAGAATAAAAACTCCTGCATAAACGATTTACAGGAATTTGTATTGCAATTTCTTA  
AGATGAAAGGAACAGCCACCAAGCAGTTTCACTCACTTTACTGATTTCTGTGTGGACTGAGTACATTCAGCTG  
ACGAATTTAGTTCCAGGAAGATGGATTGATGTTCACTAGCTTGGACAACCTCTGCAAAATATGAGACTATTTCC  
ACTTGGGAAAAATTACAACAGCAAAAAAAAAAAAAAAAAAAAAA

111/615

**FIGURE 110**

MYLSRSLSIHALWVTVSSVMQPYPLVWGHYDLCKTQIYTEEGKVWDYMACQPESTDMTKYLKVKLDPPDITCGDP  
PETFCAMGNPYMCNNECDASTPELAHPPELMFDFEGRHPSTFWQSATWKEYPKPLQVNITLSWSKTIELTDNIVI  
TFESGRPDQMLEKSLDYGRTWQPYQYYATDCLDAFHMDPKSVKDLSQHTVLEIICTEEYSTGYTTNSKIIHFEI  
KDRFALFAGPRLRNMASLYGQLDTTKKLRDFFVTDLRIRLLRPAVGEIFVDELHLARYFYAISDIKVRGRCKCN  
LHATVCVYDNSKLTCECEHNTTGPDCGKCKKNYQGRPWSPGSYLPIPKGTANTCIPSISSIGTNVCDNELLHCQN  
GGTCHNNVRCLCPAAYTGILCEKLRCEEAGSCGSDSGQGAPPHGTPALLLLTTLTGASPLVF

112/615

**FIGURE 111**

CGGTGCCGTCAGCTCGCCGGGCACCGCGGCCTCGCCCTCGCCCTCCGCCCTGCGCCTGCACCGCGTAGACCGAC  
CCCCCCTCCAGCGCGCCACCCGGTAGAGGACCCCCCGCCGTGCCCGGACCGGTCCCCGCCTTTTGTAAAACT  
TAAAGCGGGCGCAGCATTAAACGCTTCCCGCCCCGGTGACCTCTCAGGGGTCTCCCCGCCAAAGGTGCTCCGCCGC  
TAAGGAACATGGCGAAGGTGGAGCAGGTCTGAGCCTCGAGCCGCGAGCAGAGCTCAAATTCGAGGTCCCTTCA  
CCGATGTTGTCAACCACCAACCTAAAGCTTGGCAACCCGACAGACCGAAATGTGTGTTTTAAGGTGAAGACTACAG  
CACCACGTAGGTACTGTGTGAGGCCCAACAGCGGAATCATCGATGCAGGGGCTCAATTAATGTATCTGTGATGT  
TACAGCCTTTTCGATTATGATCCCAATGAGAAAAGTAAACACAAGTTTATGGTTCAGTCTATGTTTGCTCCAACTG  
ACACTTCAGATATGGAAGCAGTATGGAAGGAGGCAAAACCGGAAGACCTTATGGATTCAAACTTAGATGTGTGT  
TTGAATTGCCAGCAGAGAATGATAAACACATGATGTAGAAATAAATAAAATTATATCCACAACATGCATCAAAGA  
CAGAAACACCAATAGTGTCTAAGTCTCTGAGTTCTTCTTTGGATGACACCGAAGTTAAGAAGGTTATGGAAGAAT  
GTAAGAGGCTGCAAGGTGAAGTTCAGAGGCTACGGGAGGAGAACAGCAGTTCAAGGAAGAAGATGGACTGCGGA  
TGAGGAAGACAGTGCAGAGCAACAGCCCCATTTCAGCATTAGCCCCAACTGGGAAGGAAGAAGGCCTTAGCACCC  
GGCTCTTGGCTCTGGTGGTTTTGTTCTTTATCGTTGGTGAATTATTGGGAAGATTGCCTTGATAGAGGTAGCATG  
CACAGGATGGTAAATTGGATTGGTGGATCCACCATATCATGGGATTTAAATTTATCATAACCATGTGTAAAAAGA  
AATTAATGTATGATGACATCTCACAGGTCTTGCCCTTTAAATTACCCCTCCCTGCACACACATACACAGATACACA  
CACACAAATATAATGTAACGATCTTTAGAAAGTTAAAAATGTATAGTAACTGATTGAGGGGAAAAAGAATGAT  
CTTTTATTAAATGACAAGGGAAACCATGAGTAATGCCACAATGGCATATTGTAAATGTCATTTTAAACATTGGTAGG  
CCTTGGTACATGATGCTGGATTACCTCTCTTAAATGACACCCCTTCTCGCCTGTTGGTGTGGCCCTTGGGGAG  
CTGGAGCCCAGCATGCTGGGGAGTGCGGTGAGCTCCACACAGTAGTCCCCACGTGGCCCACTCCCGGCCAGGCT  
GCTTTCCTGTCTTTCAGTTCGTGTCCAAGCCATCAGCTCCTTGGGACTGATGAACAGAGTCAGAAGCCCAAAGGAA  
TTGCACTGTGGCAGCATCAGACGTAATCGTCATAAGTGAGAGGCGTGTGTGACTGATTGACCCAGCGCTTTGGA  
AATAAATGGCAGTGCTTTGTTCACTTAAAGGGACCAAGCTAAATTTGTATTGGTTCATGTAGTGAAGTCAAACCTG  
TTATTCAGAGATGTTTAAATGCATATTTAACTTATTTAATGTATTTTCATCTCATGTTTTCTTATTGTCACAAGAGT  
ACAGTTAATGCTGCGTGCTGCTGAACTCTGTTGGGTGAACTGGTATTGCTGCTGGAGGGCTGTGGGCTCCTCTGT  
CTCTGGAGAGTCTGGTTCATGTGGAGGTGGGGTTTATTGGGATGCTGGAGAAGAGCTGCCAGGAAGTGTTTTTCT  
GGGTGAGTAAATAACAACCTGTCTAGGGAGGGAAATTTCTAGTAGTGACAGTCAACTCTAGGTTACCTTTTTTAA  
TGAAGAGTAGTCAGTCTTCTAGATTGTTCTTATACCACCTCTCAACCATTACTCACACTCCAGCGCCAGGTCC  
AAGTCTGAGCCTGACCTCCCCTTGGGGACCTAGCCTGGAGTCAGGACAAATGGATCGGGCTGCAGAGGGTTAGAA  
GCGAGGGCACACAGCAGTTGTGGTGGGGAGCAAGGAAGAGAGAACTCTTCAGCGAATCCTTCTAGTACTAGTT  
GAGAGTTTGAAGTGAATTAATTTTATGCCATAAAAGACCAACCCAGTTCTGTTTGACTATGTAGCATCTTGAAA  
AGAAAAATTATAATAAAGCCCCAAATTAAGAAAA

113/615

**FIGURE 112**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA53977

&lt;subunit 1 of 1, 243 aa, 1 stop

&lt;MW: 27228, pI: 7.43, NX(S/T): 2

MAKVEQVLSLEPQHELKFRGPFDTDVVTNLKLGNPTRDNVCFKVKTTAPRRYCVRPNNGIIDAGASINVSVMLOP  
FDYDPNEKSKHKFMVQSMFAPDTSDMEAVWKEAKPEDLMDSKLRCVFELPAENDKPHDVEINKIISTASKTET  
PIVSKSLSSSLDDTEVKKVMEECKRLQGEVQRLREENKQFKEEDGLRMRKTVQSNPISALAPTGKEEGLSTRLL  
ALVVLFFIVGVIIGKIAL**Important features:****Transmembrane domain:**

amino acids 224-239

**N-glycosylation site.**

amino acids 68-71

**N-myristoylation site.**

amino acids 59-64, 64-69 and 235-240

114/615

**FIGURE 113**

CCCACGCGTCCGGGTGACCTGGGCCGAGCCCTCCCGGTCCGGCTAAGATTGCTGAGGAGGCGGCGGGTAGCTGGCA  
GGCGCCGACTTCCGAAGGCCGCCGTCCGGGCGAGGTGTCCTCATGACTTCTCTTGTGGACC**ATGT**CCGTGATCTT  
TTTTGCCTGCGTGGTACGGGTAAGGGATGGACTGCCCTCTCAGCCTCTACTGATTTTTACCACACCCAAGATTT  
TTTGGAATGGAGGAGACGGCTCAAGAGTTTAGCCTTGCGACTGGCCAGTATCCAGGTGAGGTCTGCAGAAGG  
TTGTGACTTTAGTATACATTTTTCTTCTTTCCGGGACGTGGCCTGCATGGCTATCTGCTCCTGCCAGTGTCCAGC  
AGCCATGGCCTTCTGCTTCCTGGAGACCCTGTGGTGGGAATTCACAGCTTCCTATGACACTACCTGCATTGGCCT  
AGCCTCCAGGCCATACGCTTTTCTTGAGTTTGACAGCATCATTCAGAAAGTGAAGTGGCATTTTAACTATGTAAG  
TTCCTCTCAGATGGAGTGCAGCTTGAAAAAATTCAGGAGGAGCTCAAGTTGCAGCCTCCAGCGGTTCTCACTCT  
GGAGGACACAGATGTGGCAAATGGGGTGATGAATGGTCACACACCGATGCACTTGGAGCCTGCTCCTAATTTCCG  
AATGGAACCACTGACAGCCCTGGGTATCCTCTCCCTCATTCTCAACATCATGTGTGCTGCCCTGAATCTCATTCTG  
AGGAGTTCACCTTGCAGAACATTCTTTACAGGATCCAAGGAGCTGGTCTGCTGGTTGGACCAAACCTCG**TGAGC**  
CAGCCACCCCTGACCCAAATGAGGAGAGCTCTGATTCTCCCATCCGGGAGCAGTGATGTCAAACCTTCTGCTGCTG  
GGGAAATCTCATCAGCAGGGAGCCTGTGGAAAAGGGCATGTCAGTGAAATCTGGGAATGGCTGGATTCCGGAAACA  
TCTGCCCATGTGTATTGATGGCAGAGCTGTGCCCACAAGCGCCTTTTATTTAGGGTAAATTAACAAATCCATT  
CTATTCCTCTGACCCATGCTTAGTACATATGACCTTTAACCCTTACATTTATATGATTCTGGGGTTGCTTCAGAA  
GTGTTATTTTCATGAATCATTCATATGATTTGATCCCCCAGGATTCTATTTTGTTTAATGGGCTTTTCTACTAAAA  
GCATAAAATACTGAGGCTGATTTAGTCAGGGCAAACCATTTACTTTACATATTCGTTTTCAATACTTGCTGTTC  
ATGTTACACAAGCTTCTTACGGTTTTCTTGTAACAATAAATATTTTGAGTAAATAATGGGTACATTTTAACAAAC  
TCAGTAGTACAACCTAACTTGTATAAAAGTGTGTAAAAATGTATAGCCATTTATATCCTATGTATAAATTAAT  
GAGGTGGCTTCAGAAATGGCAGAATAAATCTAAAGTGTTTATTAAAAAAAAAAAAAAAAAAAAAAAG

115/615

**FIGURE 114**

MSVIFACVVRVDGLPLSASTDFYHTQDFLEWRRRLKSLALRLAQYPGRGSAEGCDFSIHFSSFGDVACMAICS  
CQCPAAMAFCFLETLWWEFTASYDTTCIGLASRPYAFLEFDSIIQKVWHFNYVSSSQMECSLEKIQEELKLQPP  
AVLTLEDTDVANGVMNGHTPMHLEPAPNFRMEPVTALGILSLILNIMCAALNLIRGVHLAEHSLQDPRSWFCWLDQTS

116/615

**FIGURE 115**

CTCAGCGGGCGCTTCCTCGTAGCGAGCCTAGTGGCGGGTGTTCGATTGAAACGTGAGCGCGACCCGACCTTAAAG  
AGTGGGGAGCAAAGGGAGGACAGAGCCCTTTAAAACGAGGCGGGTGGTGCCTGCCCCCTTTAAGGGCGGGCGTCC  
GGACGACTGTATCTGAGCCCCAGACTGCCCCGAGTTTCTGTGCGCAGGCTGCGAGGAAAGGCCCTAGGCTGGGTC  
TGGGTGCTTGGCGGGCGGGCGCTTCCTCCCCGCTCGTCTCCCCGGGGCCAGAGGCACCTCGGCTTCAGTCATGCT  
GAGCAGAGTATGGAAGCACCTGACTACGAAGTGCTATCCGTGCGAGAACAGCTATTCCACGAGAGGATCCGCGAG  
TGTATTATATCAACACTTCTGTTTGCAACACTGTACATCCTCTGCCACATCTTCTGACCCGCTTCAAGAAGCCT  
GCTGAGTTCACCACAGTGGATGATGAAGATGCCACCGTCAACAAGATTGCGCTCGAGCTGTGCACCTTTACCCCTG  
GCAATTGCCCTGGGTGCTGTCTGCTCCTGCCCTTCTCCATCATCAGCAATGAGGTGCTGCTCTCCCTGCCCTCGG  
AACTACTACATCCAGTGGCTCAACGGCTCCCTCATCCATGGCCTCTGGAACCTTGTTCCTTCCCCAACCTG  
TCCCTCATCTTCTCATGCCCTTTGCATATTTCTTCACTGAGTCTGAGGGCTTTGCTGGCTCCAGAAAGGGTGTG  
CTGGGCGGGTCTATGAGACAGTGGTGATGTTGATGCTCCTCACTCTGCTGGTGCTAGGTATGGTGTGGGTGGCA  
TCAGCCATTGTGGACAAGAACAAGGCCAACAGAGAGTCACTCTATGACTTTTGGGAGTACTATCTCCCTACCTC  
TACTCATGCATCTCCTTCTTGGGGTCTGCTGCTCCTGGTGTGTAAGTCCACTGGGTCTCGCCCGCATGTTCTCC  
GTCACCTGGGAAGCTGCTAGTCAAGCCCCGGCTGCTGGAAGACCTGGAGGAGCAGCTGTACTGCTCAGCCTTTGAG  
GAGGCAGCCCTGACCCGCGAGGATCTGTAATCCTACTTCTGCTGGCTGCCCTTAGACATGGAGCTGCTACACAGA  
CAGGTCTGGCTCTGCAGACACAGAGGGTCTGCTGGAGAAGAGGCGGAAGGCTTCAGCCTGGCAACGGAACCTG  
GGCTACCCCTGGCTATGCTGTGCTTGTGCTGGTGTGACGGGCTGTCTGTGCTCATTGTGGCCATCCACATCCTG  
GAGCTGCTCATCGATGAGGCTGCCATGCCCCGAGGCATGCAGGGTACCTCCTTAGGCCAGGTCTCCTTCTCCAAG  
CTGGGCTCCTTTGGTGCCGTCAATCAGGTTGTACTCATCTTTTACCTAATGGTGTCTCAGTTGTGGGCTTCTAT  
AGCTCTCCACTCTTCCGGAGCCTGCGGCCCCAGATGGCACGACACTGCCATGACGCAGATAATTGGGAAGTGTGTC  
TGTCTCCTGGTCTAAGCTCAGCACTTCTGTCTTCTCTCGAACCCTGGGGCTCACTCGCTTTGACCTGCTGGGT  
GACTTTGGACGCTTCAACTGGCTGGGCAATTTCTACATTGTGTTCTCTACAACGCAGCCTTTGCAGGCTCACC  
ACACTCTGTCTGGTGAAGACCTTCACTGCAGCTGTGCGGGCAGAGCTGATCCGGGCTTTGGGCTGGACAGACTG  
CCGCTGCCCCGTCTCCGGTTTCCCCCAGGCATCTAGGAAGACCCAGCACCAGTGAACCTCCAGCTGGGGGTGGGAAG  
GAAAAAAGTGGACACTGCCATCTGCTGCCTAGGCCTGGAGGGAAGCCCAAGGCTACTTGGACCTCAGGACCTGGA  
ATCTGAGAGGGTGGGTGGCAGAGGGGAGCAGAGCCATCTGCACTATTGCATAATCTGAGCCAGAGTTTGGGACCA  
GGACCTCCTGCTTTTCCATACTTAAGTGTGGCCTCAGCATGGGGTAGGGCTGGGTGACTGGGTCTAGCCCCCTGAT  
CCCAAATCTGTTTACACATCAATCTGCCTCACTGCTGTTCTGGGCCATCCCCATAGCCATGTTTACATGATTGA  
TGTGCAATAGGGTGGGGTAGGGGCAGGGAAGGACTGGGCCAGGGCAGGCTCGGGAGATAGATTGTCTCCCTTGC  
CTCTGGCCCAGCAGAGCCTAAGCACTGTGCTATCCTGGAGGGGCTTTGGACCACCTGAAAGACCAAGGGGATAGG  
GAGGAGGAGGCTTCAGCCATCAGCAATAAAGTTGATCCCAGGGAAAAAA

117/615

**FIGURE 116**

MEAPDYEVLSVREQLFHERIRECIISTLLFATLYILCHIFLTRFKKPAEFTTVDDDEDATVNKIALELCTFTLAIA  
LGAVLLLPFSIISNEVLLSLPRNYIQWLNGSLIHGLWNLVLFPNLSLIFLMPFAYFFTESEGFAGSRKGVLGR  
VYETVVMMLLLTLLVLGMVWVASAIVDKNKANRESLYDFWEYYPYLYSCISFLGVLLLLVCTPLGLARMFSVTG  
KLLVKPRILLEDLEEQLYCSAFEEAALTRRICNPTSCWLPLDMELLHRQVLALQTRVLLKRRKASAWQRNLGYP  
LAMLCLLVLTGLSVLIVAIHILELLIDEAAMPGRMQGTSLGQVSFSKLGSFGAVIQVVLIFYLMVSSVVGIFYSSP  
LFRSLRPRWHDAMTQIIIGNCVCLLVLSALPVFSRTLGLTRFDLLGDFGRENWLGNFYIVFLYNAAFAGLTTLC  
LVKTFTA AVRAELIRAFGLDRLPLPVSGFPQASRKTQHQ



118/615

**FIGURE 117**

[illegible]

119/615

**FIGURE 118**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA54002
><subunit 1 of 1, 544 aa, 1. stop
><MW: 60268, pI: 9.53, NX(S/T): 3
MLLPLLLSSLLGGSQAMDGRFWIRVQESVMVPEGLCISVPCSFSPYPRQDWTGSTPAYGYWFKAVTETTKGAPVAT
NHQSREVEVEMSTRGRFQLTGDPKAGNCSLVIRDAQMQDESQYFFRVERGSYVTYNFMNDGFFLKVTVLSFTPRPQD
HNTDLTCHVDFSRKGVSAQRTVRLRVAYAPRDLVISISRDNTPALEPQPQGNVPYLEAQKGQFLRLLCAADSQPP
ATLSWVLQNRVLSSSHPWGPRPLGLELPGVKAGDSGRYTCRAENRLGSQQRALDLSVQYPPENLRVMVSQANRTV
LENLGNGTSLPVLEGQSLCLVCVTHSSPPARLSWTQRGQVLSPSQPSDPGVLELPRVQVEHEGEFTCHARHPLGS
QHVSLSLSVHYKKGLISTAFSNGAFLGIGITALLFLCLALIIMKILPKRRTQTETPRPRFSRHSTILDYINVVPT
AGPLAQKRNQKATPNSPRTPPPPGAPSPESKKNQKKQYQLPSFPEPKSSTQAPESQESQEELHYATLNFPGVRPR
PEARMPKGTQADYAEVKFQ
```

**Important features:****Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 399-418

**N-glycosylation site.**

amino acids 100-103, 297-300 and 306-309

**Immunoglobulins and major histocompatibility complex proteins signature.**

amino acids 365-371

120/615

**FIGURE 119**

CTCGCGCAGGGATCGTCCC**ATG**CGCGGGGCTCGGAGCCGCGACCCCTTGGGGGGCCTCCGGGATTTGCTACCTTTT  
TGGCTCCCTGCTCGTGAAGTCTCTTCTCACGGGCTGTGCGCTTCAATCTGGACGTGATGGGTGCCTTGCSCAA  
GGAGGGCGAGCCAGGCAGCCTCTTCGGCTTCTCTGTGGCCCTGCACCGGCAGTTGCAGCCCCGACCCAGAGCTG  
GCTGCTGGTGGGTGCTCCCCAGGCCCTGGCTCTTCTGGGCAGCAGGCGAATCGCACTGGAGGCCTCTTCGCTTG  
CCCGTTGAGCCTGGAGGAGACTGACTGCTACAGAGTGGACATCGACCAGGGAGCTGATATGCAAAAGGAAAGCAA  
GGAGAACCAGTGGTTGGGAGTCAGTGTTTCGGAGCCAGGGGCTGGGGGCAAGATTGTTACCTGTGCACACCGATA  
TGAGGCAAGGCAGCGAGTGGACCAGATCCTGGAGACGCGGGATATGATTGGTCGCTGCTTTGTGCTCAGCCAGGA  
CCTGGCCATCCGGGATGAGTTGGATGGTGGGGAATGGAAGTTCTGTGAGGGACGCCCCAAGGCCATGAACAATT  
TGGGTTCTGCCAGCAGGGCACAGCTGCCGCCTTCTCCCTGATAGCCACTACCTCCTCTTTGGGGCCCCAGGAAC  
CTATAATTGGAAGGGCACGGCCAGGGTGGAGCTCTGTGCACAGGGCTCAGCGGACCTGGCACACCTGGACGACGG  
TCCCTACGAGGGCGGGGAGAGAAGGAGCAGGACCCCGCCTCATCCCGTCCCTGCCAACAGCTACTTTGGCTT  
CTCTATTGACTCGGGGAAAGGTCTGGTGCGTGCAGAAAGCATGAGCTTTGTGGCTGGAGCCCCCGCGCCAAACCA  
CAAGGGTGTGTGGTTCATCTGCGCAAGGACAGCGCCAGTGCCTGGTGGCCGAGGTTATGCTGTCTGGGGAGCG  
CCTGACCTCCGGCTTTGGCTACTCACTGGCTGTGGCTGACCTCAACAGTGTGGCTGGCCAGACCTGATAGTGGG  
TGCCCCCTACTTCTTTGAGCGCCAAGAAGAGCTGGGGGGTGTGTGTATGTGTACTTGAACCAGGGGGGTCACTG  
GGCTGGGATCTCCCTCTCCGGCTCTGCGGCTCCCTGACTCCATGTTCCGGATCAGCCTGGCTGTCTGGGGGA  
CCTCAACCAAGATGGCTTTCCAGATATTGCAGTGGGTGCCCCCTTTGATGGTGATGGGAAAGTCTTCATCTACCA  
TGGGAGCAGCCTGGGGGTTGTGCCAAACCTTCACAGGTGCTGGAGGGCGAGGCTGTGGGCATCAAGAGCTTCGG  
CTACTCCCTGTGAGGCAGCTTGGATATGGATGGGAACCAATACCCTGACCTGCTGGTGGGCTCCCTGGCTGACAC  
CGCAGTGCTCTTCAGGGCCAGCCCATCCTCCATGTCTCCATGAGGTCTCTATTGCTCCACGAAGCATCGACCT  
GGAGCAGCCCAACTGTGCTGGCGGCCACTCGGTCTGTGTGGACCTAAGGGTCTGTTTCAGCTACATTGCAGTCCC  
CAGCAGCTATAGCCCTACTGTGGCCCTGGACTATGTGTTAGATGCGGACACAGACCGGAGGCTCCGGGGCCAGGT  
TCCCCGTGTGACCTTCTGAGCCGTAACCTGGAAGAACCCAAAGCACCAGGCCTCGGGCACCCGTGTGGCTGAAGCA  
CCAGCATGACCGAGTCTGTGGAGACGCCATGTTCCAGCTCCAGGAAAATGTCAAAGACAAGCTTCGGGGCCATTGT  
AGTGACCTTGTCTACAGTCTCCAGACCCCTCGGCTCCGGCGACAGGCTCCTGGCCAGGGGCTGCCTCCAGTGGC  
CCCCATCCTCAATGCCACCAGCCAGCACCAGCGGGCAGAGATCCACTTCTGAAGCAAGGCTGTGGTGAAGA  
CAAGATCTGCCAGAGCAATCTGCAGCTGGTCCACGCCCGCTTCTGTACCCGGGTGAGCGACACGGAATTCCAACC  
TCTGCCCCATGGATGTGGATGGAACAACAGCCCTGTTTGACTGAGTGGGCAGCCAGTCATTGGCCCTGGAGCTGAT  
GGTCACCAACCTGCCATCGGACCCAGCCAGCCCGCCAGGCTGATGGGGATGATGCCCATGAAGCCAGCTCCTGGT  
CATGCTTCTGACTCACTGCACTACTCAGGGGTCCGGGCCCTGGACCCCTGCGGAGAAGCCACTCTGCCTGTCCAA  
TGAGAATGCCTCCCATGTTGAGTGTGAGCTGGGGAACCCCATGAAGAGAGGTGCCAGGTCACCTTCTACCTCAT  
CCTTAGCACCTCCGGGATCAGCATTGAGACCACGGAAGTGGAGGTAGAGCTGCTGTTGGCCACGATCAGTGAGCA  
GGAGCTGCATCCAGTCTCTGCACGAGCCCGTGTCTTATTGAGCTGCCACTGTCCATTGCAGGAATGGCCATTCC  
CCCAACTCTTCTTCTCTGGTGTGGTGGGGGCGAGAGAGCCATGCACTGTGAGCGGGATGTGGGCAGCAAGGT  
CAAGTATGAGGTACCGGTTTCCAACCAAGGCCAGTCGCTCAGAACCCTGGGCTCTGCCTTCTCAACATCATGTG  
GCCTCATGAGATTGCCAATGGGAAGTGGTTGCTGTACCCAATGCAGGTTGAGCTGGAGGGCGGGCAGGGGCCCTGG  
GCAGAAAGGGCTTTGCTCTCCAGGCCCAACATCCTCCACCTGGATGTGGACAGTAGGGATAGGAGGCGCGGGGA  
GCTGGAGCCACCTGAGCAGCAGGAGCCTGGTGGCGGCAGGAGCCAGCATGTCTGGTGGCCAGTGTCTCTGC  
TGAGAAGAAGAAAAACATCACCTGGACTGCGCCCGGGGCGACGGCCAACTGTGTGGTGTTCAGCTGCCCACTCTA  
CAGCTTTGACCGCGCGGCTGTGCTGCATGTCTGGGGCCGTCTCTGGAACAGCACCTTTCTGGAGGAGTACTCAGC  
TGTGAAGTCCCTGGAAGTGATTGTCCGGGCCAACATCACAGTGAAGTCTCCATAAAGAACTTGATGCTCCGAGA  
TGCTTCCACAGTGATCCAGTGATGGTATACTTGGACCCCATGGCTGTGGTGGCAGAAGGAGTGGCTGGTGGGT  
CATCTCCTGGCTGTACTGGCTGGGCTGCTGGTGTAGCACTGCTGGTGTGCTCCTGTGGAAGATGGGATTCTT  
CAAACGGGCGAAGCACCCCGAGGCCACCGTGCCCCAGTACCATGCGGTGAAGATTCTCGGGAAGACCGACAGCA  
GTTCAAGGAGGAGAAGACGGGCACCATCCTGAGGAACAAGTGGGGCAGCCCCGGCGGGAGGGGCCGGATGCACA  
CCCCATCCTGGCTGTGACGGGCATCCCGAGCTGGGCCCCGATGGGCATCCAGGGCCAGGCACCGCC**TAG**GTTCC  
CATGTCCCAGCCTGGCCTGTGGCTGCCCTCCATCCCTTCCCCAGAGATGGCTCCTTGGGATGAAGAGGGTAGAGT  
GGGCTGCTGGTGTGCGCATCAAGATTTGGCAGGATCGGCTTCTCAGGGGCACAGACCTCTCCACCCACAAGAAC  
TCCTCCCACCAACTTCCCCTTAGAGTGTGTGAGATGAGAGTGGGTAAATCAGGGACAGGGCCATGGGGTAGGG  
TGAGAAGGGCAGGGGTGTCTGATGCAAGGTGGGGAGAAGGGATCCTAATCCCTTCTCTCCCATTCACCCCTGT  
GTAACAGGACCCCAAGGACCTGCCTCCCCGAAGTGCCTTAACCTAGAGGGTGGGGAGGAGGTTGTGTCACTGA  
CTCAGGCTGCTCCTTCTCTAGTTTCCCTCTCATTTGACCTTAGTTTGCTGCCATCAGTCTAGTGGTTTCGTGGT  
TTCGTCTATTATTAAAAATATTTGAGAAACAAAAA

121/615

**FIGURE 120**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA55737
><subunit 1 of 1, 1141 aa, 1 stop
><MW: 124671, pI: 5.82, NX(S/T): 5
MAGARSRDPWGASGICYLFGSLLVELLFSRAVAFNLDVMGALRKEGEPGSLFGFSVALHRQLQPRPQSWLLVGAP
QALALPGQQANRTGGLFACPLSLEETDCYRVDIDQGADMQKESKENQWLGVSVRSQGPGGKIVTCAHRYEARQRV
DQILETRDMIGRCFVLSQDLAIRDELDDGGGEWKFCGRPQGHEQFGFCQQGTAAAFSPDSHYLLFGAPGTYNWKGT
ARVELCAQGSADLAHLDDGFPYEAGGEKEQDPRILPVPANSYFGFSIDSGKGLVRAEELS FVAGAPRANHKGAVVI
LRKDSASRLVPEVMLSGERLTSFGYSLAVADLNSDGWPD LIVGAPYFFERQEELGGAVYVYLNQGGHWAGISPL
RLCGSPDSMFGISLAVLGDNLNQDGFDPDIAGVAPFDGDGKVFYIHGSSLGVVAKPSQVLEGEAVGIKSFYSLSGS
LDMDGNQYPDLLVGLADTAVLFRARPILHVSHEVSIAPRSIDLEQPN CAGGHSVCVDLRVCFSYIAVPSSYSPT
VALDYVLADTDRRLRGQVPRVTFLSRNLEEPKHQASGTVWLKHQHDRVCGDAMFQLQENVKDKLRAIVVTL SYS
LQTPRLRRQAPGQGLPPVAPILNAHQPSQRAEIHFLKQGCGEDKICQSNLQLVHARFCTRVSDTEFQPLPMDVD
GTTALFALSGQPVIGLELMVTNLPSPDPAQPQADGDDAHEAQLLVMLPDSLHYSGVRALDPAEKPLCLSNENASHV
ECELGNPMKRG AQVTFYLLISTSGISIETTELEVELELLLATISEQELHPVSARARVFIELPLSIAGMAIPQQLFFS
GVVRGERAMQSERDVGSKVKEYEVTVSNQGSRLTLGSAFLNIMWPHEIANGKWL LYPMQVELEGGQGPQKGLCS
PRPNILHLDVDSRDRRRRELEPPEQEPGERQEPMSWWPVSSAEKKKNITLDCARGTANCVVFSCPLYSFDRAA
VLHVWGR LWNSTFLEEYS AVKSLEVIVRANITVKSSIKNLMLRDASTVIPVMVYLDPM AVVAEGVPWWVILLAVL
AGLLVLALLVLLLWKMGGFFKRAKHPEATVPQYHAVKIPREDRQQFK EECTGTILRNNWGSPRREGPD AHPILAAD
GHPELGPDGHPGPGTA
```

**Important features:****Signal peptide:**

amino acids 1-33

**Transmembrane domain:**

amino acids 1040-1062

**N-glycosylation sites.**

amino acids 86-89, 746-749, 949-952, 985-988 and 1005-1008

**Integrins alpha chain proteins.**

amino acids 1064-1071, 384-408, 1041-1071, 317-346, 443-465, 385-407, 215-224, 634-647, 85-99, 322-346, 470-479, 442-466, 379-408 and 1031-1047

122/615

**FIGURE 121**

GGCACGAGGCGGCGGGGCGAGTCGCGGGATGCGCCCGGGAGCCACAGCCTGAGGCCCTCAGGTCTCTGCAGGTGTC  
GTGGAGGAACCTAGCACCTGCCATCCTCTTCCCCAATTTGCCACTTCCAGCAGCTTTAGCCCATGAGGAGGATGT  
GACCGGGAGTGAAGTCAGGAGCCCTCTGGAAGCATGGAGACTGTGGTGATTGTTGCCATAGGTGTGCTGGCCACCA  
TCTTTCTGGCTTCGTTTGCAGCCTTGGTGCTGGTTTGCAGGCAGCGCTACTGCCGGCCGCGAGACCTGCTGCAGC  
GCTATGATTCTAAGCCCATTTGTGGACCTCATTGGTGCCATGGAGACCCAGTCTGAGCCCTCTGAGTTAGAAGTGG  
ACGATGTCGTTATCACCAACCCCCACATTGAGGCCATTCTGGAGAATGAAGACTGGATCGAAGATGCCTCGGGTC  
TCATGTCCCACTGCATTGCCATCTTGAAGATTTGTCACACTCTGACAGAGAAGCTTGTGGCATGACAATGGGCT  
CTGGGGCCAAGATGAAGACTTCAGCCAGTGTGAGCGACATCATTGTGGTGGCCAAGCGGATCAGCCCCAGGGTGG  
ATGATGTTGTGAAGTCGATGTACCCTCCGTTGGACCCCCAACTCCTGGACGCACGGACGACTGCCCTGCTCCTGT  
CTGTCAGTCACCTGGTGCTGGTGACAAGGAATGCCTGCCATCTGACGGGAGGCCTGGACTGGATTGACCAGTCTC  
TGTCGGCTGCTGAGGAGCATTTGGAAGTCCTTCGAGAAGCAGCCCTAGCTTCTGAGCCAGATAAAGGCCTCCAG  
GCCCTGAAGGCTTCCTGCAGGAGCAGTCTGCAATTTAGTGCCTACAGGCCAGCAGCTAGCCATGAAGGCCCTGC  
CGCCATCCCTGGATGGCTCAGCTTAGCCTTCTACTTTTTCTATAGAGTTAGTTGTTCTCCACGGCTGGAGAGTT  
CAGCTGTGTGTGCATAGTAAAGCAGGAGATCCCCGTGAGTTTATGCCTCTTTTGCAGTTGCAAACCTGTGGCTGGT  
GAGTGGCAGTCTAATACTACAGTTAGGGGAGATGCCATTCACTCTCTGCAAGAGGAGTATTGAAAACCTGGTGGAC  
TGTCAGCTTTATTTAGCTCACCTAGTGTTCAGAAAATTGAGCCACCGTCTAAGAAATCAAGAGGTTTCACAT  
TAAAATTAGAATTTCTGGCCTCTCTCGATCGGTGAGAATGTGTGGCAATTCTGATCTGCATTTTCAGAAGAGGAC  
AATCAATTGAACTAAGTAGGGGTTTCTTCTTTTGGCAAGACTTGTAATCTCTCACCTGGCCTGTTTCATTTATT  
TGTATTATCTGCCTGGTCCCTGAGGCGTCTGGGTCTCTCCTCTCCCTTGCAAGGTTTGGGTTTGAAGCTGAGGAAC  
TACAAAGTTGATGATTTCTTTTTTATCTTTATGCCTGCAATTTTACCTAGCTACCACTAGGTGGATAGTAAATTT  
ATACTTATGTTTCCCTCAAAAAAAAAAAAAA

123/615

**FIGURE 122**

METVVIVAIGVLATIFLASFAALVIVCRQRYCRPRDLLQRYDSKPIVDLIGAMETQSEPSELEDDVVITNPHIE  
AILENEDWIEDASGLMSHCIAILKICHTLTEKLVAMTMGSGAKMKTSASVSDIIVVAKRISPRVDDVVKSMYPPL  
DPKLLDARTTALLSVSHLVLVTRNACHLTGGLDWIDQSLSAEEHLEVLREAALASEPDKGLPGPEGFLQEQSAI

124/615

**FIGURE 123**

CCCTTACATCCTCCTAGGACCCGGTCGGTAGTCGTCGCCCCAGCCCGCCGGGGGCGCAGCGCCCGAGCCGCGGCC  
CTCGAGACGGGACCGAGAGCATCATGGGCAGCACTGTCCCGCGCTCCGCCTCCGTGCTGCTTCTGCTGCTGCTCC  
TGCGCCGGGCGGAGCAGCCCTGCGGGGCGGAGCTCACCTTCGAGCTGCCGGACAACGCCAAGCAGTGCTTCCACG  
AGGAGGTGGAGCAGGGCGTGAAGTTCTCCCTGGATTACCAGGTCATCACTGGAGGCCACTACGATGTTGACTGCT  
ATGTAGAGGACCCCCAGGGGAACACCATCTACAGAGAAACGAAGAAGCAGTACGACAGCTTCACGTACCGGGCTG  
AAGTCAAGGGCGTTTATCAGTTTTGCTTCAGTAATGAGTTTTCCACCTTCTCTCACAAGACCGTCTACTTTGACT  
TTCAAGTGGGCGATGAGCCTCCCATTTCTCCAGACATGGGGAACAGGGTCACAGCTCTCACCCAGATGGAGTCCG  
CCTGCGTGACCATCCATGAGGCTCTGAAAACGGTGATTGACTCCCAGACGCATTACCGGCTGCGGGAGGCCCAGG  
ACCGGGCCCCGAGCGGAAGACCTTAATAGCCGAGTCTCTTACTGGTCTGTTGGCGAGACGATTGCCCTGTTCTGTTG  
TCAGCTTCAGTCAGTGCTACTGTTGAAAAGCTTCTTCACAGAAAACGACCCATCAGCAGGGCAGTCCACTCCT  
AGCCCCGGCATCCTGCTCTAGGGCCCCCTCATGCCCCAGGCTGGAGCAGCTCTCCTAGGTACAGCCTGCTGGGCT  
GGGTGCGTAGCCAGGGTGGAGGCAGAACGATGCTGCTGTGGTAGCCCTTTGCCTTTTCATGCCCATGCTTGATT  
CTTGACCTCAGCAGCTGAAGTCTCAGAGACCAGTAATCAGAAGGCATCCGACTGCATTAAGTGTGCAGCGCTG  
AAAAGACATTTACAACCTAGGCCAGGGATTAGCCACTGTGGGAGGCTGGACAGGCAATGGTTCAGTGGCCTGGCTG  
TTGGCAGGAACCTCCAAGTGCCCAGGCCTCTTGGGCAGCTTAGGGCCCTGCCTCTGTTTCATGATGCATGGGTCT  
TTGTCTTGGGTGTCTATCCCATATGGAGAAGAAAGGGGCTCTAAGTTCTGGCTCTTCTTTCTTTGGGGTTCTCT  
GTACCTGAGGAAACCAGGCCCTGGGTGACTTTGCAGATCTGCTCACCTCGGTGAGCAACAGTGTGAGCCATGCA  
AGCAGGACAGAATGGTGACTGGGTGCCCTTGGTGAGCTGTGTATTTCTAGGAGGTAGAAAAGTGTGGGAACTG  
TGGCTAATAAAAACTAAGTGTGAGCGTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

125/615

**FIGURE 124**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56052
<subunit 1 of 1, 217 aa, 1 stop
<MW: 24777, pI: 5.55, NX(S/T): 0
MGSTVPRSASVLLLLLLLLRRAEQPCGAELTFELPDNAKQCFHEEVEQGVKFSLDYQVITG
GHYDVDCYVEDPQGNTIYRETKKQYDSFTYRAEVKGVYQFCFSNEFSTFSHKTVYFDFQV
GDEPPILPDMGNRVTALTQMESACVTIHEALKTVIDSQTHYRLREAQDRARAEDLNSRVS
YWSVGETIALFVVSFSQVLLKSFTEKRPISRAVHS
```

**Important features:****Signal peptide:**

amino acids: 1-23

**Transmembrane domain:**

amino acids: 187-201

**N-myristoylation sites:**

amino acids: 26-32, 48-54, 131-137

**Tyrosine kinase phosphorylation site:**

amino acids: 82-91

**Glycosyl hydrolases family 25 proteins:**

amino acids: 53-61



126/615

**FIGURE 125**

GGCACGAGGCGCTGTCCACCCGGGGGCGTGGGAGTGAGGTACCAGATTAGCCCATTTGGCCCCGACGCCTCTGT  
TCTCGGAATCCGGGTGCTGCGGATTGAGGTCCCGGTTCCTAACGGACTGCAAGATGGAGGAAGGCGGGAACCTAG  
GAGGCCTGATTAAGATGGTCCATCTACTGGTCTTGTGAGGTGCCTGGGGCATGCAAATGTGGGTGACCTTCGTCT  
CAGGCTTCCTGCTTTTCCGAAGCCTTCCCCGACATACCTTCGGACTAGTGCAGAGCAAACCTCTCCCTTCTACT  
TCCACATCTCCATGGGCTGTGCCTTCATCAACCTCTGCATCTTGGCTTCACAGCATGCTTGGGCTCAGCTCACAT  
TCTGGGAGGCCAGCCAGCTTTACCTGCTGTTCCCTGAGCCTTACGCTGGCCACTGTCAACGCCCGCTGGCTGGAAC  
CCCGCACCACAGCTGCCATGTGGGCCCTGCAAACCGTGGAGAAGGAGCGAGGCCTGGGTGGGAGGTACCAGGCA  
GCCACCAGGGTCCCGATCCCTACCGCCAGCTGCGAGAGAAGGACCCCAAGTACAGTGCTCTCCGCCAGAATTTCT  
TCCGCTACCATGGGCTGTCTCTCTTTGCAATCTGGGCTGCGTCCTGAGCAATGGGCTCTGTCTCGCTGGCCTTG  
CCCTGGAAATAAGGAGCCTCTAGCATGGGCCCTGCATGCTAATAAATGCTTCTTCAGAAATGAAAAAAAAAAAAA  
AAAAAA

127/615

**FIGURE 126**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56107

&lt;subunit 1 of 1, 231 aa, 1 stop

&lt;NX(S/T): 0

MEEGGNLGGLIKVMVHLLVLSGAWGMQMWVTFVSGFLLFRSLPRHTFGLVQSKLFPFYFHISMGCAFINLCILASQ  
HAWAQLTFWEASQLYLLFLSITLATVNARWLEPRTTAAMWALQTVKEKERGLGGEVPGSHQGPDPYRQLREKDPKY  
SALRQNFRRYHGLSSLCNLGCVLSNGLCLAGLALEIRSL

**Signal peptide:**

amino acids 1-24

**Transmembrane domain:**

amino acids 86-103, 60-75

**Casein kinase II phosphorylation site.**

amino acids 82-86

**Tyrosine kinase phosphorylation site.**

amino acids 144-151

**N-myristoylation site.**

amino acids 4-10, 5-11, 47-53, 170-176, 176-182

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 54-65

**G-protein coupled receptors proteins.**

amino acids 44-85

128/615

**FIGURE 127**

GCTTCATTTCTCCCGACTCAGCTTCCCACCCTGGGCTTTCCGAGGTGCTTTTCGCCGCTGTCCCCACCACTGCAGC  
CATGATCTCCTTAACGGACACGCAGAAAATTGGAATGGGATTAACAGGATTTGGAGTGTTTTTCTGTCTTTGG  
AATGATTCTCTTTTTTGACAAAGCACTACTGGCTATTGGAAATGTTTTATTTGTAGCCGGCTTGGCTTTGTAAT  
TGGTTTAGAAAGAACATTCAGATTCTTCTTCCAAAAACATAAAATGAAAGCTACAGGTTTTTTCTGGGTGGTGT  
ATTTGTAGTCCTTATTGGTTGGCCTTTGATAGGCATGATCTTCGAAATTTATGGATTTTTCTCTGTTCAGGGG  
CTTCTTTCCTGTCGTTGTTGGCTTTATTAGAAGAGTGCCAGTCCTTGGATCCCTCCTAAATTTACCTGGAATTAG  
ATCATTGTAGATAAAGTTGGAGAAAGCAACAATATGGTATTAACAACAAGTGAATTTGAAGACTCATTTAAAATA  
TTGTGTTATTTATAAAGTCATTTGAAGAATATTGAGCACAAAAATTAAATTACATGAAATAGCTTGAATGTTCTT  
TACAGGAGTTTAAAACGTATAGCCTACAAAGTACCAGCAGCAAATTAGCAAAGAAGCAGTGAAAACAGGCTTCTA  
CTCAAGTGAAGTAAGAAGAAGTCAGCAAGCAAAGTGAAGAGAGTGAATCCATGTTAATGATGCTTAAGAAACTC  
TTGAAGGCTATTTGTGTTGTTTTCCACAATGTGCGAAACTCAGCCATCCTTAGAGAACTGTGGTGCCTGTTCT  
TTTCTTTTTATTTTGAAGGCTCAGGAGCATCCATAGGCATTTGCTTTTTAGAAGTGTCCACTGCAATGGCAAAAA  
TATTTCCAGTTGCACTGTATCTCTGGAAGTGATGCATGAATTCGATTGGATTGTGTCATTTTAAAGTATTAAAC  
CAAGGAAACCCAATTTTGATGTATGGATTACTTTTTTTTGNCCNCAGGGCC

129/615

## **FIGURE 128**

MISLTDQKIGMGLTGFGVFFLEFFGMILFFDKALLAIGNVLFVAGLAFVIGLERTFRFFFQKHMKATGFFLGGV  
FVVLIGWPLIGMIFEIYGFFLLFRGFFPVVVGFIIRVPVLGSLNLPGIRSFVDKVGESNNMV

**Important features:**

**Transmembrane domains:**

amino acids 12-30 (typeII), 33-52, 69-89 and 93-109

**N-myristoylation sites.**

amino acids 11-16, 51-56 and 116-121

**Aminoacyl-transfer RNA synthetases class-II protein.**

amino acids 49-59

130/615

**FIGURE 129**

AATTCAGATTTTAAAGCCCATCTGCAGTGGAATTTTCATGAAGTAGCAAGAGGACACCATCTTCTTGATTATACA  
AGAAAGGAGTGTACCTATCACACACAGGGGGAAAAATGCTCTTTTGGGTGCTAGGCCTCCTAATCCTCTGTGGTT  
TTCTGTGGACTCGTAAAGGAAAACTAAAGATTGAAGACATCACTGATAAGTACATTTTTATCACTGGATGTGACT  
CGGGCTTTGAAACTTGGCAGCCAGAACTTTTGATAAAAAGGGATTTTCATGTAATCGCTGCCTGTCTGACTGAAT  
CAGGATCAACAGCTTTAAAGGCAGAAACCTCAGAGAGACTTCGTACTGTGCTTCTGGATGTGACCGACCCAGAGA  
ATGTCAAGAGGACTGCCCAGTGGGTGAAGAACCAAGTTGGGGAGAAAGGTCTCTGGGGTCTGATCAATAATGCTG  
GTGTTCCCGGCGTGCTGGCTCCCACTGACTGGCTGACACTAGAGGACTACAGAGAACCTATTGAAGTGAACCTGT  
TTGGACTCATCAGTGTGACACTAAATATGCTTCCTTTGGTCAAGAAAGCTCAAGGGAGAGTTATTAATGTCTCCA  
GTGTTGGAGGTCGCCTTGCAATCGTTGGAGGGGGCTATACTCCATCCAAATATGCAGTGGAAGGTTTCAATGACA  
GCTTAAGACGGGACATGAAAGCTTTTGGTGTGCACGTCTCATGCATTGAACCAGGATTGTTCAAAACAACTTGG  
CAGATCCAGTAAAGGTAATTGAAAAAACTCGCCATTGGGAGCAGCTGTCTCCAGACATCAAACAACAATATG  
GAGAAGGTTACATTGAAAAAGTCTAGACAACTGAAAGGCAATAAATCCTATGTGAACATGGACCTCTCTCCGG  
TGGTAGAGTGCATGGACCACGCTCTAACAAGTCTCTTCCCTAAGACTCATTATGCCGCTGGAAAAGATGCCAAAA  
TTTTCTGGATACCTCTGTCTCACATGCCAGCAGCTTTGCAAGACTTTTTATTGTTGAAACAGAAAGCAGAGCTGG  
CTAATCCCAAGGCAGTGTGACTCAGCTAACCACAAATGTCTCCTCCAGGCTATGAAATTGGCCGATTTCAAGAAC  
ACATCTCCTTTTCAACCCCATTCCTTATCTGTCTCAACCTGGACTCATTTAGATCGTGCTTATTTGGATTGCAAA  
AGGGAGTCCCACCATCGCTGGTGGTATCCCAGGGTCCCTGCTCAAGTTTTCTTTGAAAGGAGGGCTGGAATGGT  
ACATCACATAGGCAAGTCCTGCCCTGTATTTAGGCTTTGCCTGCTTGGTGTGATGTAAGGGAAATTGAAAGACTT  
GCCCATTCAAAATGATCTTTACCGTGGCTGCCCATGCTTATGGTCCCCAGCATTTACAGTAACTTGTGAATGT  
TAAGTATCATCTCTTATCTAAATATTTAAAGATAAGTCAACCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAA

131/615

**FIGURE 130**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56406
><subunit 1 of 1, 319 aa, 1 stop
><MW: 35227, pI: 8.97, NX(S/T): 3
MLFWVLGLLILCGFLWTRKGLKIEDITDKYIFITGCDSGFGNLAARTFDKKGFHVIAACLTESGSTALKAETSE
RLRTVLLDVTDPENVKRTAQWVKNQVGEKGLWGLINNAGVPGVLAPTDWLTLEDYREPIEVNLFGLISVTLNMLP
LVKKAQGRVINVSSVGGRLAIVGGGYTPSKYAVEGFNDLRRDMKAFGVHVSIEPGLFKTNLADPVKVEKKLA
IWEQLSPDIKQQYGEGYIEKSLDKLKGKNSYVNMDLSPVVECDHALTSLFPKTHYAAGKDAKIFWIPLSHMPAA
LQDFLLLKQKAE LANPKAV
```

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**Transmembrane domain:**

amino acids 136-152

**N-glycosylation sites.**

amino acids 161-163, 187-190 and 253-256

**Glycosaminoglycan attachment site.**

amino acids 39-42

**N-myristoylation sites.**

amino acids 36-41, 42-47, 108-113, 166-171, 198-203 and 207-212

132/615

**FIGURE 131**

AGACAGTACCTCCTCCCTAGGACTACACAAGGACTGAACCAGAAGGAAGAGGACAGAGCAAAGCCATGAACATCA  
TCCTAGAAATCCTTCTGCTTCTGATCACCATCATCTACTCCTACTTGGAGTCGTTGGTGAAGTTTTTCATTCCTC  
AGAGGAGAAAATCTGTGGCTGGGGAGATTGTTCTCATTACTGGAGCTGGGCATGGAATAGGCAGGCAGACTACTT  
ATGAATTTGCAAAACGACAGAGCATATTGGTTCTGTGGGATATTAATAAGCGCGGTGTGGAGGAACTGCAGCTG  
AGTGCCGAAAACCTAGGCGTCACTGCGCATGCGTATGTGGTAGACTGCAGCAACAGAGAAGAGATCTATCGCTCTC  
TAAATCAGGTGAAGAAAGAAGTGGGTGATGTAACAATCGTGGTGAATAATGCTGGGACAGTATATCCAGCCGATC  
TTCTCAGCACCAAGGATGAAGAGATTACCAAGACATTTGAGGTCAACATCCTAGGACATTTTTGGATCACAAAAG  
CACTTCTTCCATCGATGATGGAGAGAAATCATGGCCACATCGTCACAGTGGCTTCAGTGTGCGGCCACGAAGGGA  
TTCCTTACCTCATCCCATATTGTTCCAGCAAATTTGCCGCTGTTGGCTTTCACAGAGGTCTGACATCAGAACTTC  
AGGCCTTGGGAAAAAAGTGGTATCAAACCTCATGTCTCTGCCAGTTTTTGTGAATACTGGGTTACCAAAAAATC  
CAAGCACAAAGATTATGGCCTGTATTGGAGACAGATGAAGTCGTAAGAAGTCTGATAGATGGAATACTTACCAATA  
AGAAAATGATTTTTGTTCCATCGTATATCAATATCTTCTGAGACTACAGAAGTTTCTTCCTGAACGCGCCTCAG  
CGATTTTAAATCGTATGCAGAATATTCAATTTGAAGCAGTGGTTGGCCACAAAATCAAATGAAATTGAATAAATA  
AGCTCCAGCCAGAGATGTATGCATGATAATGATATGAATAGTTTCGAATCAATGCTGCAAAGCTTTATTTACAT  
TTTTTCAGTCCTGATAATATTA AAAACATTGGTTTGGCACTAGCAGCAGTCAAACGAACAAGATTAATTACCTGT  
CTTCCTGTTTCTCAAGAATATTTACGTAGTTTTTCATAGGTCTGTTTTCTTCCTTTCATGCCTCTTAAAAACTTCTG  
TGCTTACATAAAACATACTTAAAAGGTTTTCTTTAAGATATTTTATTTTTCCATTTAAAGGTGGACAAAAGCTACC  
TCCCTAAAAGTAAATACAAAGAGAACTTATTTACACAGGGAAGGTTTAAGACTGTTCAAGTAGCATTCCAATCTG  
TAGCCATGCCACAGAATATCAACAAGAACACAGAATGAGTGCACAGCTAAGAGATCAAGTTTCAGCAGGCAGCTT  
TATCTCAACCTGGACATATTTTAAGATTCAGCATTGAAAGATTTCCCTAGCCTCTTCCTTTTTTCATTAGCCCAA  
AACGGTGCAACTCTATTCTGGACTTTATTACTTGATTCTGTCTTCTGTATAACTCTGAAGTCCACAAAAGTGGA  
CCCTCTATATTTCTCCCTTTTTTATAGTCTTATAAGATACATTATGAAAGGTGACCGACTCTATTTTAAATCTCA  
GAATTTTAAGTTCTAGCCCCATGATAACCTTTTTCTTTGTAATTTATGCTTTCATATATCCTTGGTCCCAGAGAT  
GTTTAGACAATTTTAGGCTCAAAAATTAAAGCTAACACAGGAAAAGGAAGTGTACTGGCTATTACATAAGAAACA  
ATGGACCCAAGAGAAGAA

133/615

**FIGURE 132**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56409

&lt;subunit 1 of 1, 300 aa, 1 stop

&lt;MW: 33655, pI: 9.31, NX(S/T): 1

MNIILEILLLLITIIYSYLESIVKFFIPQRRKSVAGEIVLITGAGHGIGRQTTYEFAKRQSILVLWDINKRGVEE  
TAAECRKLGVTAHAYVVDCSNREEIYRSLNQVKEVGDTVIVVNNAGTVYPADLLSTKDEEITKTFEVLNIGHFW  
ITKALLPSMMERNHGHIVTVASVCGHEGIPYLI PYCSSKFAAVGFHRGLTSELQALGKTGIKTSCLCPVFNVTGF  
TKNPSTRLWPVLETDEVVRSLIDGILTNNKMIFVPSYINIFLRLQKFLPERASAILNRMQNIQFEAVVGHKIKMK

**Important features:****Signal peptide:**

amino acids 1-19

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 30-33 and 58-61

**Short-chain alcohol dehydrogenase family protein**

amino acids 165-202, 37-49, 112-122 and 210-219



134/615

**FIGURE 133**

CTGAGGCGGCGGTAGC**ATG**GAGGGGGAGAGTACGTGCGCGGTGCTCTCGGGCTTTGTGCTCGGCGCACTCGCTTT  
CCAGCACCTCAACACGGACTCGGACACGGAAGGTTTTCTTCTGGGGAAGTAAAAGGTGAAGCCAAGAACAGCAT  
TACTGATTCCCAAATGGATGATGTTGAAGTTGTTTATACAATTGACATTCAGAAATATATTCCATGCTATCAGCT  
TTTTAGCTTTTATAATTCTTCAGGCGAAGTAAATGAGCAAGCACTGAAGAAAATATTATCAAATGTCAAAAAGAA  
TGTGGTAGGTTGGTACAAATTCCGTCGTCAATTCAGATCAGATCATGACGTTTAGAGAGAGGCTGCTTCACAAAAA  
CTTGCAGGAGCATTTTTCAAACCAAGACCTTGTTTTCTGCTATTAACACCAAGTATAATAACAGAAAGCTGCTC  
TACTCATCGACTGGAACATTCCTTATATAAACCTCAAAAAGGACTTTTTTCACAGGGTACCTTTAGTGGTTGCCAA  
TCTGGGCATGTCTGAACAACCTGGGTTATAAACTGTATCAGGTTCTGTATGTCCACTGGTTTTAGCCGAGCAGT  
ACAAACACACAGCTCTAAATTTTTGAAGAAGATGGATCCTTAAAGGAGGTACATAAGATAAATGAAATGTATGC  
TTCATTACAAGAGGAATTAAAGAGTATATGCAAAAAAGTGGAAGACAGTGAACAAGCAGTAGATAAACTAGTAAA  
GGATGTAAACAGATTAAAAACGAGAAATTGAGAAAAGGAGAGGAGCACAGATTCAAGGCAGCAAGAGAGAAGAACAT  
CCAAAAAGACCCTCAGGAGAACATTTTTCTTTGTCAGGCATTACGGACCTTTTTTCCAAATTCTGAATTTCTTCA  
TTCATGTGTTATGTCTTTAAAAAATAGACATGTTTCTAAAAGTAGCTGTAACACACCACCATCTCGATGTAGT  
AGACAATCTGACCTTAATGGTAGAACACACTGACATTCCTGAAGCTAGTCCAGCTAGTACACCACAAATCATTA  
GCATAAGCCTTAGACTTAGATGACAGATGGCAATTCAGAGATCTCGGTTGTTAGATACACAAGACAAACGATC  
TAAAGCAAATACTGGTAGTAGTAACCAAGATAAAGCATCCAAAATGAGCAGCCCAGAAACAGATGAAGAAATTGA  
AAAGATGAAGGGTTTTGGTGAATATTCACGGTCTCCTACATTT**TGA**TCTTTTAACTTACAAGGAGATTTTTT  
ATTTGGCTGATGGGTAAAGCCAAACATTTCTATTGTTTTACTATGTTGAGCTACTTGACAGTAAGTTCATTTGTT  
TTTACTATGTTACCTGTTTGCAGTAATACACAGATAACTCTTAGTGCATTTACTTCACAAAGTACTTTTTCAA  
CATCAGATGCTTTTATTTCCAAACCTTTTTTTCACCTTTCACTAAGTTGTTGAGGGGAAGGCTTACACAGACACA  
TTCTTTAGAATTGGAAAAGTGAGACCAGGCACAGTGGCTCACACCTGTAATCCCAGCACTTAGGGAAGACAAGTC  
AGGAGGATTGATTGAAGCTAGGAGTTAGAGACCAGCCTGGGCAACGTATTGAGACCATGTCTATTAAAAAATAAA  
ATGGAAAAGCAAGAATAGCCTTATTTTCAAAATATGGAAAGAAATTTATATGAAAATTTATCTGAGTCATTAAAA  
TTCTCCTTAAGTGATACTTTTTTAGAAGTACATTATGGCTAGAGTTGCCAGATAAAATGCTGGATATCATGCAAT  
AAATTTGCAAAACATCATCTAAAATTTAAAAA

135/615

**FIGURE 134**

MEGESTSAVLSGFVLGALAFQHLNTDSDTEGFLGGEVKGEAKNSITDSQMDDVEVVYTTIDIQKYIPCYQLFSFY  
SSGEVNEQALKKILSNVKKNVVGWYKERRHSDQIMTFRERLHKNLQEHFSNQDLVFLLLTPSIITESCSTHRLE  
HSLYKPQKGLFHRVPLVVANLGMSEQLGYKTVSGSCMSTGFSAVQTHSSKFFEDGSLKEVHKINEMYASLQEE  
LKSICKKVEDSEQAVDKLVKDVNRLKREIEKRRGAQIQAAAREKNIQKDPQENIFLCQALRTFFPNSEFLHSCVMS  
LKNRHVSKSSCNYNHHLVDVNDLTLMVEHTDIPEASPASTPQIIKHKALDLDLDRWQFKRSRLDLDQDKRSKANTG  
SSNQDKASKMSSPETDEEIEKMKGFGEYSRSPTE

136/615

**FIGURE 135**

GGCACAGCCGCGCGGGCGGAGGGCAGAGTCAGCCGAGCCGAGTCCAGCCGGACGAGCGGACCAGCGCAGGGCAGCC  
CAAGCAGCGCGCAGCGAACGCCCCGCCGCCGCCACACCCTCTGCGGTCCCCGCGGCGCCTGCCACCCTTCCCTCC  
TTCCCCGCGTCCCCGCGCTCGCCGGCCAGTCAGCTTGCCGGGTTGCTGCCCCGCGAAACCCGAGGTCAACAGCC  
CGCGCTCTGCTTCCCTGGGCGCGCGCCGCTCCACGCCCTCCTTCTCCCTGGCCCCGGCGCCTGGCACCCGGGG  
ACCGTTGCCTGACGCGAGGCCAGCTCTACTTTTCGCCCCGCGTCTCCTCCGCTGCTCGCTCTTCCACCAACT  
CCAACCTCCTTCTCCCTCCAGCTCCACTCGCTAGTCCCCGACTCCGCCAGCCCTCGGCCGCTGCCGTAGCGCGC  
TTCCCGTCCGGTCCCAAAGGTGGGAACGCGTCCGCCCGGCCCGCACCATGGCACGGTTCGGCTTGCCCGCGCTT  
CTCTGCACCCCTGGCAGTGCTCAGCGCCGCGCTGCTGGCTGCCGAGCTCAAGTCGAAAAGTTGCTCGGAAGTGCGA  
CGTCTTTACGTGTCCAAAGGCTTCAACAAGAACGATGCCCCCTCCACGAGATCAACGGTGATCATTGAAGATC  
TGTCCCCAGGGTCTACCTGCTGCTCTCAAGAGATGGAGGAGAAGTACAGCCTGCAAAGTAAAGATGATTTCAA  
AGTGTGCTCAGCGAACAGTGCAATCATTTGCAAGCTGTCTTTGCTTACGTTACAAGAAGTTTGATGAATTCTTC  
AAAGAATACTTGAATGCAGAGAAATCCCTGAATGATATGTTTGTGAAGACATATGGCCATTTATACATGCAA  
AATTCTGAGCTATTTAAAGATCTCTTCGTAGAGTTGAAACGTTACTACGTGGTGGGAAATGTGAACCTGGAAGAA  
ATGCTAAATGACTTCTGGGCTCGCCTCCTGGAGCGGATGTTCCGCTGGTGAACCTCCAGTACCCTTTACAGAT  
GAGTATCTGGAATGTGTGAGCAAGTATACGGAGCAGCTGAAGCCCTTCGGAGATGTCCCTCGCAAATTGAAGCTC  
CAGGTTACTCGTGCTTTTGTAGCAGCCGTACTTTTCGCTCAAGGCTTAGCGGTTGCGGGAGATGTCGTGAGCAAG  
GTCTCCGTGGTAAACCCACAGCCAGTGTACCCATGCCCTGTTGAAGATGATCTACTGCTCCCACTGCCGGGGT  
CTCGTGACTGTGAAGCCATGTTACAATACTGCTCAAACATCATGAGAGGCTGTTTGGCCAACCAAGGGGATCTC  
GATTTTGAATGGAACAATTTCATAGATGCTATGCTGATGGTGGCAGAGAGGCTAGAGGGTCCCTTCAACATTGAA  
TCGGTCATGGATCCCATCGATGTGAAGATTTCTGATGCTATTATGAACATGCAGGATAATAGTGTTCAGTGTCT  
CAGAAGGTTTCCAGGGATGTGGACCCCCAAGCCCTCCAGCTGGACGAATTTCTCGTTCCATCTCTGAAAGT  
GCCTTCAGTGCTCGCTTCAGACCACATCACCCCGAGGAACGCCCAACCACAGCAGCTGGCACTAGTTTGGACCGA  
CTGGTTACTGATGTCAAGGAGAACTGAAACAGGCCAAGAAATTTGGTCTCCTCCCTCCGAGCAACGTTTGCAAC  
GATGAGAGGATGGCTGCAGGAAACGGCAATGAGGATGACTGTTGGAATGGGAAAGGCAAAAGCAGGTACCTGTTT  
GCAGTGACAGGAAATGGATTAGCCAACAGGGCAACAACCCAGAGGTCCAGGTTGACACCAGCAAACAGACATA  
CTGATCCTTCGTCAAATCATGGCTCTTCGAGTGATGACCAGCAAGATGAAGAATGCATACAATGGGAACGACGTG  
GACTTCTTTGATATCAGTGATGAAAGTAGTGGAGAAGGAAAGTGGCTGTGAGTATCAGCAGTGCCCTTCA  
GAGTTTGACTACAATGCCACTGACCATGCTGGGAAGAGTGCCAATGAGAAAGCCGACAGTGCTGGTGTCCGTCCT  
GGGGCACAGGCCTACCTCCTCACTGTCTTCTGCATCTTGTTCCTGGTTATGCAGAGAGAGTGAGATTAATTTCTCA  
AACTCTGAGAAAAAGTGTTCATCAAAAAGTTAAAAGGCACCAGTTATCACTTTTCTACCATCCTAGTGACTTTGC  
TTTTTAAATGAATGGACAACAATGTACAGTTTTTACTATGTGGCCACTGGTTTAAAGAGTGCTGACTTTGTTTC  
TCATTCACTTTGGGAGGAAAAGGACTGTGCATTGAGTTGGTTCTGCTCCCCAAACCATGTTAAACGTGGCT  
AACAGTGTAGGTACAGAACTATAGTTAGTTGTGATTTTATCACTCTATTATTTGTTTGTATGTTTTT  
TTCTCATTTTCGTTTGTGGGTTTTTTTTTCCAAGTGTGATCTCGCTTGTTCCTTACAAGCAAACAGGGTCCCTT  
CTTGGCACGTAACATGTACGTATTTCTGAAATATTAAATAGCTGTACAGAAGCAGGTTTTATTATCATGTTATC  
TTATTAAAGAAAAAGCCCAAAAGC

137/615

**FIGURE 136**

MARFGLPALLCTLAVLSAALLAAELKSKSCSEVRRLYVSKGFNKNDAPLHEINGDHLKICPQGSTCCSQEMEKEY  
SLQSKDDFKSVVSEQCNHLQAVFASRYKKFDEFFKELLENNAEKSLNDMFVKTYGHLYMQNSELFKDLFVELKRY  
VVGNNLEEMLNDFWARLLERMFLVNSQYHFTDEYLECVSKYTEQLKPFQDVPRKLLQVTRAFVAARTFAQGL  
AVAGDVVSKSVVNPTAQCTHALLKMIYCSHCRLVTVKPCYNYCSNIMRGCLANQGDLD FEWNNFIDAMLMAE  
RLEGPFNIESVMDPIDVKISDAIMNQDNSVQVSQKVFQGC GPPKPLEAGRISRSISESAF SARFRPHHPEERPT  
TAAGTSLDRLVTDVKEKLKQAKKFWSSLP SNVCNDERMAAGNGNEDDCWNGKGKSRYLFAVTGNGLANQGNNEV  
QVDTSKPDILILRQIMALRVMTSKMKNAYNGNDVDFDISDESSGEGSGSGCEYQQCPSEFDYNATDHAGKSANE  
KADSAGVRPGAQAYLLTVFCILFLVMQREWR

138/615

**FIGURE 137**

GC GGGCTGTTGACGGCGCTGCGATGGCTGCCTGCGAGGGCAGGAGAAGCGGAGCTCTCGGTTCTCTCAGTCGGA  
CTTCTGACGCCGCCAGTGGGCGGGGCCCCCTTGGGCCGTCGCCACCACTGTAGTCATGTACCCACCGCCGCCGCC  
GCCGCTCATCGGGACTTCATCTCGGTGACGCTGAGCTTTGGCGAGAGCTATGACAACAGCAAGAGTTGGCGGCG  
GCGCTCGTGTGGAGGAAATGGAAGCAACTGTCGAGATTGCAGCGGAATATGATTCTCTTCTCTCTGCTTTCT  
GCTTTTCTGTGGACTCCTCTTCTACATCAACTTGGCTGACCATTGGAAAGCTCTGGCTTTCAGGCTAGAGGAAGA  
GCAGAAGATGAGGCCAGAAATTGCTGGGTAAACCAGCAAATCCACCCGCTTACCAGCTCCTCAGAAGGCGGA  
CACCGACCTGAGAACTTACCTGAGATTTCTGCACAGAAGACACAAAGACACATCCAGCGGGGACCACCTCACCT  
GCAGATTAGACCCCAAGCCAAGACCTGAAGGATGGGACCCAGGAGGAGGCCACAAAAGGCAAGAAGCCCTGT  
GGATCCCCGCCCCGAAGGAGATCCGCAGAGGACAGTCATCAGCTGGAGGGGAGCGGTGATCGAGCCTGAGCAGGG  
CACCGAGCTCCCTTCAAGAAGAGCAGAAGTGCCACCAAGCCTCCCCTGCCACCGGCCAGGACACAGGGCACACC  
AGTGCATCTGAACATATCGCCAGAAGGGCGTGATTGACGTCTTCTGTCATGCATGGAAGGATAACCGCAAGTTTGC  
ATGGGGCCATGACGAGCTGAAGCCTGTGTCCAGGTCTTTCAGTGAGTGGTTTGGCCTCGGTCTCACACTGATCGA  
CGCGCTGGACACCATGTGGATCTTGGGTCTGAGGAAAGAATTTGAGGAAGCCAGGAAGTGGGTGTGGAAGAAGTT  
ACACTTTGAAAAGGACGTGGACGTCAACCTGTTTGAAGACGATCCGCATCCTGGGGGGGCTCCTGAGTGCCTA  
CCACCTGTCTGGGACAGCCTCTTCTGAGGAAAGCTGAGGATTTTGGAAATCGGCTAATGCCTGCCTTCAGAAC  
ACCATCCAAGATTCTTACTCGGATGTGAACATCGGTACTGGAGTTGCCACCGGCCAGGTGGACCTCCGACAG  
CACTGTGGCCGAGGTGACCAAGCTTCAAGCTGGAGTTCCGGGAGCTCTCCGCTCTCACAGGGGATAAGAAGTTTCA  
GGAGGCAGTGGAGAAGGTGACACAGCACATCCACGGCCTGTCTGGGAAGAAGGATGGGCTGGTGCCCATGTTTCA  
CAATACCCACAGTGGCCTCTTCAACCCACCTGGGCGTATTACGCTGGGCGCCAGGGCCGACAGCTACTATGAGTA  
CCTGCTGAAGCAGTGGATCCAGGGCGGGAAGCAGGAGACACAGCTGCTGGAAGACTACGTGGAAGCCATCGAGGG  
TGTCAGAACGCACCTGCTGCGGCACTCCGAGCCAGTAAGCTCACCTTTGTGGGGGAGCTTGCCACGGCCGCTT  
CAGTGCCAAGATGGACCACCTGGTGTGCTTCTGCCAGGGACGCTGGCTCTGGGCGTCTACACGGCCTGCCCGC  
CAGCCACATGGAGCTGGCCAGGAGCTCATGGAGACTTGTACCAGATGAACCGGCAGATGGAGACGGGGCTGAG  
TCCCGAGATCGTGCACTTCAACCTTACCCCCAGCCGGGCGTCCGGACGTGGAGGTCAAGCCAGCAGACAGGCA  
CAACCTGCTGCGGCCAGAGACCGTGGAGAGCCTGTTCTACCTGTACCGCGTCACAGGGGACCGCAAATACCAGGA  
CTGGGGCTGGGAGATTCTGCAGAGCTTCAAGCCGATTACACGGGTCCCCTCGGGTGGCTATTCTTCCATCAACAA  
TGTCCAGGATCCTCAGAAGCCCCGAGCCTAGGGACAAGATGGAGAGCTTCTTCTGGGGGAGACGCTCAAGTATCT  
GTTCTTGCTCTTCTCCGATGACCCAAACCTGCTCAGCCTGGACGCCTACGTGTTCAACACCGAAGCCACCTCT  
GCCTATCTGGACCCCTGCCTAGGGTGGATGGCTGCTGGTGTGGGGACTTCGGGTGGGCAGAGGCACCTTGCTGGG  
TCTGTGGCATTCTTCCAAGGGGCCACGTAGCACCGGCAACCGCCAAGTGAGCCAGGCTCTGAAGTGGCTCTGGGCT  
CCTCCTCGTCTCTGCTTTAATCAGGACACCGTGAGGACAAGTGAGGCCGTGAGTCTTGGTGTGATGCGGGGTGGG  
CTGGGCCGCTGGAGCCTCCGCCTGCTTCTCAGAAAGACACGAATCATGACTCACGATTGCTGAAGCCTGAGCAG  
GTCTCTGTGGGCCGACCAGAGGGGGGCTTCAGAGTGGTCCCTGGTACTGGGGTGACCGAGTGGACAGCCCAGGGT  
GCAGCTCTGCCCCGGGCTCGTGAAGCCTCAGATGTCCCCAATCCAAGGGTCTGGAGGGGCTGCCGTGACTCCAGAG  
GCCTGAGGCTCCAGGGCTGGCTCTGGTGTTTACAAGCTGGACTCAGGGATCCTCTGGCCGCCCGCAGGGGGCT  
TGGAGGGCTGGACGGCAAGTCCGTCTAGCTCACGGGCCCCCTCAGTGGAATGGGTCTTTTCGGTGGAGATAAAAG  
TTGATTTGCTCTAACCGCAA

139/615

**FIGURE 138**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56529
><subunit 1 of 1, 699 aa, 1 stop
><MW: 79553, pI: 7.83, NX(S/T): 0
MAACEGRRSGALGSSQSDFLTPPVGGAPWAVATTVMYPPPPPPPHRDFISVTLSTFGESYDNSKSWRRRSCWRKW
KQLSRLQRNMILFLLAFLLFCGLLFYINLADHWKALAFRLLEEQKMRPEIAGLKPANPPVLPAPQKADTDPENLP
EISSQKTQRHIQRGPPHLQIRPPSQDLKDGTQEEATKRQEAPVDPRPEGDPQRTVISWRGAVIEPEQGTELPSRR
AEVPTKPPLPPARTQGTPVHLNRYRQKGVIDVFLHAWKGYRKFAWGHDELKPVSRSEWFGGLGLTLIDALDTMWI
LGLRKEFEEARKWVSKKLHFEKDVDVNLFESTITILGGLLSAYHLSGDSLFLRKAEDFGNRLMPAFRTPSKIPYS
DVNIGTGVAHPPRWTSdstVAEVTSIQLEFRELsRLTGDKKFQEAVEKVTQHIHGLSGKKDGLVPMFINTHSGLF
THLGVFTILGARADSYEYLLKQWIIQGGKQETQLLEDYVEAIEGVrTHLLRHSEPSKLTfVgELAHGRFSakMDHL
VCFLPGTLALGVYHGLPASHMELAQELMETCYQMNRQMETGLSPEIVHFNLYPQPGRRDVEVKPADRHNLRLRPET
VESLFYLYRVTGDRKYQDWGWEILQSFSRfTRVPsGGYSSINNVQDPQKPEPRDKMESFFLGETLKYLFLLFSDD
PNLLSLDAYVFNTeAHPLPIWTPA
```

**Important features of the protein:****Transmembrane domain:**

amino acids 21-40 and 84-105 (type II)

140/615

**FIGURE 139**

CTCGCCCTCAAATGGGAACGCTGGCCTGGGACTAAAGCATAGACCACCAGGCTGAGTATCCTGACCTGAGTCATC  
CCCAGGGATCAGGAGCCTCCAGCAGGGAACCTTCCATTATATTCTTCAAGCAACTTACAGCTGCACCGACAGTTG  
CG**ATG**AAAGTTCTAATCTCTTCCCTCCTCCTGTTGCTGCCACTAATGCTGATGTCCATGGTCTCTAGCAGCCTGA  
ATCCAGGGGTGCGCCAGAGGCCACAGGGACCGAGGCCAGGCTTCTAGGAGATGGCTCCAGGAAGGCGGCCAAGAAT  
GTGAGTGCAAAGATTGGTTCTGAGAGCCCCGAGAAGAAAATTCATGACAGTGTCTGGGCTGCCAAAGAAGCAGT  
GCCCCTGTGATCATTTCAAGGGCAATGTGAAGAAAACAAGACACCAAAGGCACCACAGAAAGCCAAACAAGCATT  
CCAGAGCCTGCCAGCAATTTCTCAAACAATGTCAGCTAAGAAGCTTTGCTCTGCCTTTG**TAGG**AGCTCTGAGCGC  
CCACTCTTCCAATTAAACATTCTCAGCCAAGAAGACAGTGAGCACACCTACCAGACACTCTTCTTCTCCACCTC  
ACTCTCCCACTGTACCCACCCCTAAATCATTCCAGTGCTCTCAAAAAGCATGTTTTTCAAGATCATTTTGTGGT  
TGCTCTCTCTAGTGTCTTCTCTCTCGTCAGTCTTAGCCTGTGCCCTCCCCTTACCCAGGCTTAGGCTTAATTAC  
CTGAAAGATTCCAGGAACTGTAGCTTCTAGCTAGTGTCAATTAACCTTAAATGCAATCAGGAAAGTAGCAAAC  
AGAAGTCAATAAATATTTTAAATGTCAAAAAAAAAAAAAAAAAA

141/615

**FIGURE 140**

MKVLISLLLLLPLMLMSMVSSSLNPGVARGHRDRGQASRRWLQEGGQECECKDWFLRAPRRKFMTVSGLPKKQC  
PCDHFKGNVKKTRHQRHHRKPNKHSRACQQFLKQCQLRSFALPL



142/615

**FIGURE 141**

AATGGCTGTCTTAGTACTTCGCCTGACAGTTGTCCTGGGACTGCTTGTCTTATTCCTGACCTGCTATGCAGACGA  
CAAACCAGACAAGCCAGACGACAAGCCAGACGACTCGGGCAAAGACCCAAAGCCAGACTTCCCCAAATTCCTAAG  
CCTCCTGGGACACAGAGATCATTGAGAATGCAGTCGAGTTCATCCTCCGCTCCATGTCCAGGAGCACAGGATTTAT  
GGAATTTGATGATAATGAAGGAAAACATTCATCAAAGTGAATCCTCAGGACACACCCATGTGGCTCCTGGACAA  
TCCAAGAGCAGCCAAATCCTGCTTTTCCAGTTTGGCTCCACAAGTCTCCAGGACAGAGCCCTCAAAGCAACTCC  
CAACGAGTTCTCAGGATTCAGGCTCTGGCTTCAACCAAACAGAACTCATTTTGAACACCCTGACTGCATTTTTC  
TTTGTAGAAAGTTAGAATAAATATGGCGCTTTGGGATCACATAGTTGATGGAGAGGAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

143/615

**FIGURE 142**

MAVLVLRRLTVVLGLLVLFITCYADDKPKDPDDKPDGKDPKDPDFPKFLSLLGTEIIENAVEFILRSMSRSTGFM  
EFDDNEGKHSSK

144/615

**FIGURE 143**

GGACGCCAGCGCCTGCAGAGGCTGAGCAGGGAAAAAGCCAGTGCCCCAGCGGAAGCACAGCTCAGAGCTGGTCTG  
CC**ATG**GACATCCTGGTCCCCTCCTGCAGCTGCTGGTGCTGCTTCTTACCCTGCCCTGCACCTCATGGCTCTGC  
TGGGCTGCTGGCAGCCCCCTGTGCAAAAGCTACTTCCCCTACCTGATGGCCGTGCTGACTCCCAAGAGCAACCGCA  
AGATGGAGAGCAAGAAACGGGAGCTCTTCAGCCAGATAAAGGGGCTTACAGGAGCCTCCGGGAAAGTGGCCCTAC  
TGGAGCTGGGCTGCGGAACCGGAGCCAACTTTCAGTTCTACCCACCGGGCTGCAGGGTCACCTGCCTAGACCCAA  
ATCCCCACTTTGAGAAGTTCCTGACAAAGAGCATGGCTGAGAACAGGCACCTCCAATATGAGCGGTTTGTGGTGG  
CTCCTGGAGAGGACATGAGACAGCTGGCTGATGGCTCCATGGATGTGGTGGTCTGCACTCTGGTGCTGTGCTCTG  
TGCAGAGCCCAAGGAAGGTCTGTCAGGAGGTCCGGAGAGTACTGAGACCGGGAGGTGTGCTCTTTTTCTGGGAGC  
ATGTGGCAGAACCATATGGAAGCTGGGCCTTCATGTGGCAGCAAGTTTTCGAGCCCACCTGGAAACACATTGGGG  
ATGGCTGCTGCCTCACCAGAGAGACCTGGAAGGATCTTGAGAACGCCCAGTTCTCCGAAATCCAAATGGAACGAC  
AGCCCCCTCCCTTGAAGTGGCTACCTGTTGGGCCCCACATCATGGGAAAGGCTGTCAAACAATCTTTCCCAAGCT  
CCAAGGCACTCATTGCTCCTTCCCCAGCCTCCAATTAGAACAAGCCACCCACCAGCCTATCTATCTTCCACTGA  
GAGGGACCT**AG**CAGAATGAGAGAAGACATTGATGACACCTACTAGTCCCTCTCTCCCCAACCTCTGCCAGGGC  
AATCTCTAACTTCAATCCCGCCTTCGACAGTGAAAAAGCTCTACTTCTACGCTGACCCAGGGAGGAAACACTAGG  
ACCCTGTTGTATCCTCAACTGCAAGTTTCTGGACTAGTCTCCCAACGTTTGCCTCCCAATGTTGTCCCTTTCTCT  
CGTTCCCATGGTAAAGCTCCTCTCGCTTTCTCCTGAGGCTACACCCATGCGTCTCTAGGAACTGGTCACAAAAG  
TCATGGTGCCTGCATCCCTGCCAAGCCCCCTGACCCTCTCTCCCCACTACCACCTTCTTCTGAGCTGGGGGCA  
CCAGGGAGAATCAGAGATGCTGGGGATGCCAGAGCAAGACTCAAAGAGGCAGAGGTTTGTCTCAAATATTTT  
TAATAAATAGACGAAACCACG

145/615

**FIGURE 144**

MDILVPLLQLLVLLLTLPPLHLMALLGCWQPLCKSYFPYLMVLTTPKSNRKMESKKRELFSSQIKGLTGASGKVALL  
ELGCGTGANFQFYPPGCRVTCCLDPNPHFEKFLTKSMAENRHLQYERFVVAPGEDMRQLADGSMDVVVCTLVLCVS  
QSPRKVLQEVRRVLRPGGVLEFFWEHVAEPYGSWAFMWQQVFPTWKHIGDGCCLTRETWKDLNAQFSEIQMERQ  
PPPLKWLPGPHIMGKAVKQSFPSKALICSFPSLQLEQATHQPIYLPLRGT

146/615

**FIGURE 145**

GTGGGATTATTTGAGTGCAAGATCGTTTTCTCAGTGGTGGTGGAAAGTTGCCTCATCGCAGGCAGATGTTGGGGC  
TTTGTCCGAACAGCTCCCCTCTGCCAGCTTCTGTAGATAAGGGTTAAAACTAATATTTATATGACAGAAGAAAA  
AGATGTCATTCCGTAAAGTAAACATCATCATCTTGGTCCTGGCTGTTGCTCTCTTCTTACTGGTTTTGCACCATA  
ACTTCCTCAGCTTGAGCAGTTTGTAAAGGAATGAGGTTACAGATTGAGGAATTGTAGGGCCCAACCTATAGACT  
TTGTCCCAAATGCTCTCCGACATGCAGTAGATGGGAGACAAGAGGAGATTCCCTGTGGTCATCGCTGCATCTGAAG  
ACAGGCTTGGGGGGGCCATTGCAGCTATAAACAGCATTGAGCACAACACTCGCTCCAATGTGATTTTCTACATTG  
TTACTCTCAACAATACAGCAGACCATCTCCGGTCCTGGCTCAACAGTGATTCCCTGAAAAGCATCAGATACAAAA  
TTGTCAATTTTGACCCTAACTTTTGGGAAGGAAAAGTAAAGGAGGATCCCTGACCAGGGGGAATCCATGAAACCTT  
TAACCTTTGCAAGGTTCTACTTGCCAATTCTGGTTCCCAGCGCAAAGAAGGCCATATACATGGATGATGATGTAA  
TTGTGCAAGGTGATATTCTTGCCCTTTACAATACAGCACTGAAGCCAGGACATGCAGCTGCATTTTCAGAAGATT  
GTGATTCAGCCTCTACTAAAGTTGTTCATCCGTGGAGCAGGAAACCAGTACAATTACATTGGCTATCTTGACTATA  
AAAAGGAAAGAATTCGTAAGCTTTCCATGAAAGCCAGCACTTGCTCATTTAATCCTGGAGTTTTTGTGCAAACC  
TGACGGAATGGAAACGACAGAATATACTAACCAACTGGAAAAATGGATGAAACTCAATGTAGAAGAGGGACTGT  
ATAGCAGAACCCCTGGCTGGTAGCATCACAACACCTCCTCTGCTTATCGTATTTTATCAACAGCACTCTACCATCG  
ATCCTATGTGGAATGTCCGCCACCTTGGTTCCAGTGCTGGAAAACGATATTCACCTCAGTTTGTAAAGGCTGCCA  
AGTTACTCCATTGGAATGGACATTTGAAGCCATGGGGAAGGACTGCTTCATATACTGATGTTTGGGAAAAATGGT  
ATATTCCAGACCCCAACAGGCAATTCAACCTAATCCGAAGATATACCGAGATCTCAAACATAAAGTGAACAGAA  
TTTGAAGTGTAAAGCAAGCATTTCTCAGGAAGTCCTGGAAGATAGCATGCATGGGAAGTAACAGTTGCTAGGCTTC  
AATGCCTATCGGTAGCAAGCCATGGAAAAAGATGTGTCAGCTAGGTAAAGATGACAACTGCCCTGTCTGGCAGT  
CAGCTTCCCAGACAGACTATAGACTATAAATATGTCTCCATCTGCCTTACCAAGTGTCTTCTACTACAATGCTG  
AATGACTGGAAAGAAGAACTGATATGGCTAGTTTCAAGCTAGCTGGTACAGATAATTCAAACTGCTGTTGGTTTTA  
ATTTTGTAACTGTGGCCTGATCTGTAAATAAACTTACATTTTC

147/615

**FIGURE 146**

MSFRKVNIIILVLAVALFLLVLHHNFLSLSSLLRNEVTDSGIVGPQPIDFVPNALRHAVDGRQEEIPV VIAASED  
RLGGAIAAINSIQHNTSRNVIFYIVTLNNTADHLRSWLNSDSLKSIRYKIVNFDPKLLEGKVKEDPDQGESMKPL  
TFARFYLPILVPSAKKAIYMDDDVIVQGDILALYNTALKPGHAAAFSEDCDSASTKVVIRGAGNQYNYIGYLDYK  
KERIRKLSMKASTCSFNPGVFVANLTEWKRONITNQLEKWMKLNVEEGLYSRTL AGSITTPPLLIVFYQQHSTID  
PMWNVRLGSSAGKRYSPQFVKAALLHWNGHLKPWGRTASYTDVWEKWIIPDPTGKFNLIRRYTEISNIK

148/615

**FIGURE 147**

GTTTGAATTCCTTCAACTATACCCACAGTCCAAAAGCAGACTCACTGTGTCCCAGGCTACCAGTTCCTCCAAGCA  
AGTCATTTCCCTTATTTAACCGATGTGTCCCTCAAACACCTGAGTGCTACTCCCTATTTGCATCTGTTTTGATAA  
ATGATGTTGACACCCTCCACCGAATTCTAAGTGGAAATCATGTCGGGAAGAGATACAATCCTTGGCCTGTGTATCC  
TCGCATTAGCCTTGTCTTTGGCCATGATGTTTACCTTCAGATTCAACACCCTTCTGGTTCACATTTTCATTT  
CATTGGTTATTTTGGGATTGTTGTTTGTCTGCGGTGTTTTATGGTGGCTGTATTATGACTATACCAACGACCTCA  
GCATAGAATTGGACACAGAAAGGGAAAATATGAAGTGCCTGCTGGGGTTTGCTATCGTATCCACAGGCATCACGG  
CAGTGCTGCTCGTCTTGATTTTTGTTCTCAGAAAGAGAATAAAATTGACAGTTGAGCTTTTCCAAATCACAAATA  
AAGCCATCAGCAGTGCTCCCTTCCTGCTGTTCCAGCCACTGTGGACATTTGCCATCCTCATTTTCTTCTGGGTCC  
TCTGGGTGGCTGTGCTGCTGAGCCTGGGAAGTGCAGGAGCTGCCCAGGTTATGGAAGGCGGCCAAGTGAATATA  
AGCCCCTTTCGGGCATTTCGGTACATGTGGTCGTACCATTTAATTGGCCTCATCTGGACTAGTGAATTCATCCTTG  
CGTGCCAGCAAATGACTATAGCTGGGGCAGTGGTTACTTGTTATTTCAACAGAAGTAAAAATGATCCTCCTGATC  
ATCCCATCCTTTTCGTCTCTCTCCATTCTCTTCTTCTACCATCAAGGAACCGTTGTGAAAGGGTCATTTTTAATCT  
CTGTGGTGAGGATTCCGAGAATCATTGTCTGTACATGCAAACGCACTGAAAGAACAGCAGCATGGTGCATTGT  
CCAGGTACCTGTTCCGATGCTGCTACTGCTGTTTCTGGTGTCTTGACAAATACCTGCTCCATCTCAACCAGAATG  
CATATACTACAACGCTATTAATGGGACAGATTTCTGTACATCAGCAAAAGATGCATTCAAAATCTTGTCCAAGA  
ACTCAAGTCACCTTACATCTATTAAGTGGTCTTTGGAGACTTCATAATTTTCTAGGAAAGGTGTTAGTGGTGTGTT  
TCACTGTTTTTGGAGGACTCATGGCTTTTAACTACAATCGGGCATTCAGGTGTGGGCAGTCCCTCTGTTATTGG  
TAGCTTTTTTTGCCTACTTAGTAGCCCATAGTTTTTTATCTGTGTTTGAAACTGTGCTGGATGCACCTTTTCTGT  
GTTTTGCTGTTGATCTGGAAACAAATGATGGATCGTCAGAAAAGCCCTACTTTATGGATCAAGAATTTCTGAGTT  
TCGTAAAAAGGAGCAACAAATTAACAATGCAAGGGCACAGCAGGACAAGCACTCATTAAAGGAATGAGGAGGGAA  
CAGAACTCCAGGCCATTGTGAGATAGATACCCATTTAGGTATCTGTACCTGGAAAACATTTCTTCTAAGAGCCA  
TTTACAGAATAGAAGATGAGACCACTAGAGAAAAGTTAGTGAATTTTTTTTTAAAGACCTAATAAACCCCTATTC  
TTCCTCAAAA

149/615

**FIGURE 148**

MSGRDTILGLCIIALALSLAMMFTFRFITTLVHIFISLVILGLLFVCGVLWWLYDYTDNLSIELDTERENMKC  
VLGFAIVSTGITAVLLVLIFVLRKRIKLTVELFQITNKAISSAPFLLFQPLWTFAILIFFWVLWVAVLLSLGTA  
AAQVMEGGQVEYKPLSGIRYMWSYHLIGLIWTSEFILACQQMTIAGAVVTCYFNRSKNDPPDHPILSSLSILFFY  
HQGTVVKGSFLISVVRIPRIIVMYMQNALKEQQHGALSRYLFRCCYCCFWCLDKYLLHLNQNAYTTTAINGTDFC  
TSAKDAFKILSKNSSHFTSINCFGDFIIFLGKVLVVCFTVFGGMLMAFNYNRAFQVWAVPLLLVAFFAYLVHSFL  
SVFETVLDALFLCFAVDLETNDGSSEKPYFMDQEFLSFVKRSNKLNNARAQQDKHSLRNEEGTELQAIVR



**FIGURE 149**

GTTCGATTAGCTCCTCTGAGAAGAAGAGAAAAAGGTTCTTGACCTCTCCCTGTTTCTTCCTTAGAATAATTTGTA  
 TGGGATTTGTGATGCAGGAAGAAGCTTAAGGGAATAAATAATTCATTCTGTGTGGTGAANAATTTTTTGAAAAAAA  
 ATTGCCCTTCTTCAAACAAGGTTGTCATTCTGATATTTATGAGGACTGTTGTTCTTCACTATGAAGGCATCTGTTAT  
 TGAAATGTTCCTTGTTTTGCTGGTGACTGGAGTACATTCAAACAAAGAAACGGCAAAGAAGATTAAAGGCCCAA  
 GTTCACTGTGCCCTCAGATCAACTGCGATGTCAAAGCCGGAAGATCATCGATCTTGAGTTTCATTGTGAATGTCC  
 AGCAGGATGCCAAGACCCCAATAACCATGTTATTGGCATGACGTGATGCATCCTACTCCAGTGTGTGTGGCG  
 TGCCGTGACACAGTGGTGTGCTTGATAATTCAGGAGGGAAATACTTGTTCGGAAGGTTGCTGGACAGTCTGGTTA  
 CAAAGGGAGCTTATCCAACGGTGCCCAATCGTTATCCCTACCACGATGGAGAGAATCCTTTATCGTCTTAGAAAG  
 TAAACCCAAAAGGGGTGAACCTTACCATCAGCTCTTACATATCATCATCGAAAGTCCAGCTGCCAAGCAGG  
 TGAGACCACAAAAGCCTATCAGAGGCCACCTATTCCAGGGACAATCGACACAGCCGGTCACTCTGATGCAGCTTCT  
 GGCTGTCACTGTAGCTGTGGCCACCCACCACCTTGCCAAAGGCCATCCCTTCTGCTGCTTCTACCACAGCAT  
 CCCCAGACCACAATCAGTGGGCGACAGGAGGCCAGGAGATGGATCTCTGGTCCACTGCCACCTACACAAGCAGCCA  
 GAACAGGCCCCAGAGCTGATCCAGGTATCAAAGGCAAGATCCTTCAGGAGCTGCCTCCAGAAACCTGTTGGAGC  
 GGATGTCACTGGCTGGGACTTGTTCGTAAGAAGAATTGAGCACACAGTCTTTGGAGCCGATTCCTGGGAGATCC  
 AAACCTGCAAAATTGACTTGTGCTTTTTAATTGATGGGAGCACCAGCATTTGGCAAACGGCGATTCCGAATCCAGAA  
 CGAGCTCTGGCTGATGTTGCCAACGCTCTTGACATTTGGCCCTGCCGTTCCACTGATGGTGTGTCCAGTATGG  
 AGACAACCTTGCTACTCACTTTAACTCAAGACACACAGCAATTCTCGAGATCTGAAGCAGCCATAGAAAAAT  
 TACTCAGAGAGGAGGACTTTCTAATGTAGGTGCGGCCATCTCCTTTGTGACCAAGAAGATTCTTTTCCAAAGCCAA  
 TGGAAACAGAAGAGCGGGGCTCCCAATGTGGTGGTGTATGTGGTATGGCTGCGCCACGCAGCAAAAGTGGAGGAGC  
 TTCAGACTTGCAGAGAGTCAAGGAATCAACATTTCTTCTCATCACCATTGAAGTGTCTGTAAGATGAGAAGCA  
 GTATGTGGTGGAGCCCCAATTGCAAAACAAGGCCGTGTGCAGAACAAACGGCTTCTACTCGCTCCAGTGCAGAG  
 CTGGTTTGGCCTCCACAAGACCCCTGCAGCCTCTGGTGAAGCGGGTCTGCGACACTGACCGCTGGCCTGCAGCA  
 GCTGCTTTGAACCTCGGCTGACATTTGGCTTCTGTCAAGCAGGCTCCAGCAGTGTGGGACGGGCAACTTCGCA  
 CGTCCCTCCAGTTTGTGACCAACCTACCCAAAGATTTTGAGATTTCCGACACGGACACGCGCATCGGGGCCGTGCA  
 GTACACTACGAACACGCGCTGGAGTTTGGTTTCGACAAGTACAGCAGCAAGCCTGACATCCTCAAGCCATCAA  
 GAGGGTGGGCTACTGGAGTGGTGGCACCAGCAGCGGGGCTGCCATCAACTTCGCCCTGGAGAGCTCTTCAAGAA  
 GTCCAAGGCCAACAAAGAGAAGTTAATGATCCTCATCACCAGCGGAGGTTCTACGACGACGCTCCGATCCCAGC  
 CATGGCTGCCCATCTGAAGGGAGTGAATCACTTATCGCATAGGCTGTGGCTGGGCTGCCAAGAGGAGCTAGAAGT  
 CATTGCCACTCACCCCGCCAGAGACCCTCTCTTCTTTGTGGACGAGTTTGACAACCTCCTCAGTATGTCCCCAG  
 GATCATCCAGAACATTTGTACAGAGTTCAACTCACAGCCTCGGAAGCTGAATTCAGAGCAGGCAGAGCACCAGCA  
 GTGCTGCTTTTACTAAGTACGTTGTGGACACCCCAACCGTTTAATGGGCGACGCAAGTGCATCAAGTCTTTGGC  
 AGGGATGGAGAAACAAATGTCTTGTTATTATTCTTTGCCATCATGCTTTTTTCATATCCAAAACCTTGGAGTTAC  
 AAGATGATCACAAACGTATAGAATGAGCCAAAAGGCTACATCATGTTGAGGGTCTGGAGATTTTACATTTTGA  
 CAATTGTTTTCAAATAAATGTTCCGGAATACAGTGCAGCCCTTACGACAGGCTTACGTAGAGCTTTTTGTGAGATT  
 TTTAAGTTGTTATTTCTGATTTTGAACCTCTGTAACCTCAGCAAGTTTTCATTTTGTGATGACAATGTAGGAATTG  
 CTGAATTAAATGTTTGAAGGATGAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT  
 AAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAATAAAAAAT

151/615

**FIGURE 150**

MRTVVLTMKASVIEMFLVLLVTGVHSNKETAKKIKRPKFTVPQINCDVKAGKIIDPEFIVKCPAGCQDPKYHVYG  
TDVYASYSSVCGAAVHSGVLDNSGGKILVRKVAGQSGYKGSYSNGVQSLSLPRWRESFIVLESKPKKGVITYPSAL  
TYSSSKSPAAQAGETTKAYQRFFIPGTTAQPVTLMQLLAVTVAVATPTTLPRPSPSAASTTSIPRPQSVGHRSEQE  
MDLWSTATYTSSQNRPRADPGIQRQDPSGAAPQKPVGADVSLGLVPKEELSTQSLPVS LGDPNCKIDLSFLIDG  
STSIGKRRFRIQKQLLADVAQALDIGPAGPLMGVVQYGDNPATHFNLKHTNSRDLKTAIEKITQRGGLSNVGRA  
ISFVTKNFFSKANGNRSGAPNVVVVMVDGWPTDKVEEASRLARESGINIFFITIEGAAENKQYVVEPNFANKAV  
CRTNGFYSLHVQSWFGLHKTLPVLRVCDTDRACSKTCLNSADIGFVIDGSSSVGTGNFRTVLQFVTNLTKF  
EISDTDRIGAVQYTYEQRLFEFGFDKYSSKPDILNAIKRVGYWSSGTSTGAAINFALEQLFKKSKPNKRKLMILI  
TDGRSYDDVRIPAMAAHLKGVITYAIGVAWAAQEELEVIATHPARDHSEFFVDEFDNLHQYVPRIIQNICTEFNSQPRN

152/615

**FIGURE 151**

CAGGATGAACTGGTTGCAGTGGCTGCTGCTGCTGCGGGGGCGCTGAGAGGACACGAGCTCT**ATG**CCTTTCCGGCT  
GCTCATCCCGCTCGGCCCTCCTGTGCGCGCTGCTGCCTCAGCACCATGGTGCGCCAGGTCCCGACGGCTCCGCGCC  
AGATCCCGCCCACTACAGTTTTTCTCTGACTCTAATTGATGCACTGGACACCTTGCTGATTTTGGGGAATGTCTC  
AGAATCCAAAGAGTGGTTGAAGTGCTCCAGGACAGCGTGGACTTTGATATTGATGTGAACGCCTCTGTGTTGA  
AACAAACATTTCGAGTGGTAGGAGGACTCCTGTCTGCTCATCTGCTCTCCAAGAAGGCTGGGGTGAAGTAGAGGC  
TGGATGGCCCTGTTCCGGGCCTCTCCTGAGAATGGCTGAGGAGGCGGCCCGAAAACTCCTCCAGCCTTTCAGAC  
CCCCACTGGCATGCCATATGGAACAGTGAACCTACTTTCATGGCGTGAACCCAGGAGAGACCCCTGTACCTGTAC  
GGCAGGGATTGGGACCTTCATTGTTGAATTTGCCACCCTGAGCAGCCTCACTGGTGACCCGGTGTTCGAAGATGT  
GGCCAGAGTGGCTTTGATGCGCCTCTGGGAGAGCCGGTCAGATATCGGGCTGGTCGGCAACCCACATTGATGTGCT  
CACTGGCAAGTGGGTGGCCCAGGACGCAGGCATCGGGGCTGGCGTGGACTCCTACTTTGAGTACTTGGTGAAAGG  
AGCCATCCTGCTTCAGGATAAGAAGCTCATGGCCATGTTCTAGAGTATAACAAAGCCATCCGGAACCTACACCCG  
CTTCGATGACTGGTACCTGTGGGTTGAGATGTACAAGGGGACTGTGTCCATGCCAGTCTTCCAGTCCTTGGAGGC  
CTACTGGCCTGGTCTTCAGAGCCTCATTGGAGACATTGACAATGCCATGAGGACCTTCCTCAACTACTACACTGT  
ATGGAAGCAGTTTGGGGGGCTCCCGGAATTCTACAACATTCCTCAGGGATACACAGTGGAGAAGCGAGAGGGCTA  
CCCACTTCGGCCAGAACTTATTGAAAGCGCAATGTACCTCTACCGTGCCACGGGGGATCCCAACCTCCTAGAACT  
CGGAAGAGATGCTGTGGAATCCATTGAAAAATCAGCAAGGTGGAGTGCGGATTTGCAACAATCAAAGATCTGCG  
AGACCACAAGCTGGACAACCGCATGGAGTCGTTCTTCTGCGCGAGACTGTGAAATACCTCTACCTCCTGTTTGA  
CCCAACCAACTTCATCCACAACAATGGGTCCACCTTCGACGCGGTGATCACCCCTATGGGGAGTGCATCCTGGG  
GGCTGGGGGGTACATCTTCAACACAGAAGCTCACCCATCGACCTTGCCGCCCTGCACTGCTGCCAGAGGCTGAA  
GGAAGAGCAGTGGGAGGTGGAGGACTTGATGAGGGAATTCTACTCTCTCAAACGGAGCAGGTGCAAATTTTCAGAA  
AAACACTGTTAGTTCGGGGCCATGGGAACCTCCAGCAAGGCCAGGAACACTCTTCTCACCAGAAAACCATGACCA  
GGCAAGGGAGAGGAAGCCTGCCAAACAGAAGGTCCCACTTCTCAGCTGCCCCAGTCAGCCCTTACCTCCAAGTT  
GGCATTACTGGGACAGGTTTTCTAGACTCCTCA**TA**ACCACCTGGATAATTTTTTTATTTTTTATTTTTTTGAGGCT  
AAACTATAATAAATTGCTTTTGGCTATCATAAAA

153/615

**FIGURE 152**

MPFRLLIPLGLLCALLPQHHGAPGPDGSAPDPAHYSFSLTLIDALDTLLILGNVSEFQRVVEVLQDSVDFDIDVN  
ASVFETNIRVVGGLLSAHLSSKKAGVEVEAGWPCSGPLLMAEEAARKLLPAFQTPTGMPYGTVNLLHGVNPGET  
PVTCTAGIGTFIVEFATLSSLTGDPVFEDVARVALMRLWESRSDIGLVGNHIDVLTGKWVAQDAGIGAGVDSYFE  
YLVKGAILLQDKKLMAMFLEYNKAIRNYTRFDDWYLWVQMYKGTVSMFPVFQSLEAYWPGLQSLIGDIDNAMRTFL  
NYYTVWKQFGGLPEFYNIPOGYTVEKREGYPLRPELIESAMYLYRATGDPITLLELGRDAVESIEKISKVECGFAT  
IKDLRDHKLDRMESFFLAETVKYLYLLFDPTNFIHNNGSTFDAVITPYGECILGAGGYIFNTEAHPIDLAALHC  
CQRLKEEQWEVEDLMREFYSLKRSRSKFQKNTVSSGPWEPPARPGLFSPENHDQARERKPAKQKVPLLSCPSQP  
FTSKLALLGQVFLDSS

154/615

**FIGURE 153**

CGGACGCGTGGGCGGACGCGTGGGCGGACGCGTGGGTGGGAGGGGGCAGGATGGGAGGGAAAGTGAAGAAAACA  
GAAAAGGAGAGGGACAGAGGCCAGAGGACTTCTCATACTGGACAGAAACCGATCAGGC**ATG**GAACTCCCCCTTCGT  
CACTCACCTGTTCTTGCCCCCTGGTGTTCCTGACAGGTCTCTGCTCCCCCTTTAACCTGGATGAACATCACCCACG  
CCTATTCCCAGGGCCACCAGAAGCTGAATTTGGATACAGTGTCTTACAACATGTTGGGGGTGGACAGCGATGGAT  
GCTGGTGGGCGCCCCCTGGGATGGGCCTTCAGGCGACCGGAGGGGGGACGTTTATCGCTGCCCTGTAGGGGGGGC  
CCACAATGCCCCATGTGCCAAGGGCCACTTAGGTGACTACCAACTGGGAAATTATCTCATCTGCTGTGAATAT  
GCACCTGGGGATGTCTCTGTTAGAGACAGATGGTGATGGGGGATTATGGTGAGCT**TAAG**GAGAGGGTGGTGGCAG  
TGTCTCTGAAGGTCCATAAAAGAAAAAGAGAAGTGTGGTAAGGGAAATGGTCTGTGTGGAGGGGTCAAGGAGT  
TAAAAACCTAGAAAGCAAAAGGTAGGTAATGTCAAGGAGTAGTCTTCATGCCTCCTTCAACTGGGAGCATGTTT  
TGAGGGTGGCCTCCCAAGCCTGGGAGTAACTATTTCCCCATCCCAGGCCTGTGCCCTCTCTGGTCTCGTGCT  
TGTGGCAGCTCTGTCTTCAGTTCTGGGATATGTGCCCGTGTGGATGCTTCATTCCAGCCTCAGGGAAGCCTGGCA  
CCCACTGCCAACGTGAGCCAGAGGAAGGCTGAGTACTTGGTTCACAGAAGGAGATACTGGGTGGGAAAAAGATG  
GGGCAAAGCGGTATGATGCCTGGCAAAGGGCCTGCATGGCTATCCTCATTGCTACCTAATGTGCTTGCAAAGCT  
CCATGTTTCCTAACAGATTCAGACTCCTGGCCAGGTGTGGTGGCCACACCTGTAATTCTAGCACTTTGGGAGGC  
CAAGGTGGGCAGATCACTTGAGGTCAGGAGTTCAGACCAGCCTGGCCAACATGGTGAACTCCATCTCTACTAA  
AAAAAAAAAATAAAAAATTAGCTGGGTGCGCTAGTGCATGCCTGTAATCTCATCTACTCGGGAGGCTAAGACA  
GGAGACTCTCACTTCAACCCAGGAGGTGGAGGTTGCGGTGAGCCAAGATTGTGCCTCTGCACTCTAGCGTGGGTG  
ACAGAGTAAGCGAGACTCCATCTCAAAAATAATAATAATAATTCAGACTCCTTATCAGGAGTCCATGATCTG  
GCCTGGCACAGTAACTCATGCCTGTAATCCCAACATTTGGGAGGCCAACGCAGGAGGATTGCTTGAGGTCTGGA  
GGTTTGAGACCAGCCTGGGCAACATAGAAAGACCCCATCTCTAAATAAATGTTTTAAAAAT

155/615

**FIGURE 154**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57039
><subunit 1 of 1, 124 aa, 1 stop
><MW: 13352, pI: 5.99, NX(S/T): 1
MELPFVTHLFLPLVFLTGLCSFFNLDEHHPRLFPGPPEAEFGYSVLQHVGGGQRWMLVGAPWDGPGSDDRRGDVYR
CPVGGAHNAPCAKGLG DYQLGNSSHPAVNMHLGMSLLET DGDGGFMVS
```

**Important features:****Signal peptide:**

amino acids 1-22

**Cell attachment sequence.**

amino acids 70-73

**N-glycosylation site.**

amino acids 98-101

**Integrins alpha chain proteins**

amino acids 67-81

156/615

**FIGURE 155**

GCGAGCTCCGGGTGCTGTGGCCCGGCCTTGGCGGGGCGGCCTCCGGCTCAGGCTGGCTGAGAGGCTCCCAGCTGC  
AGCGTCCCCGCCCGCCTCCTCGGGAGCTCTGATCTCAGCTGACAGTGCCCTCGGGGACCAACAAGCCTGGCAGG  
GTCTCACTTTGTTGCCAGGCTGGAGTTCAGTGCCATGATCATGGTTTACTGCAGCCTTGACCTCCTGGGTTCAA  
GCGATCCTGCTGAGTAGCTGGGACTACAGGACAAAATTAGAAGATCAAAATGGAAAAATATGCTGCTTTGGTTGAT  
ATTTTTACCCCTGGGTGGACCCCTCATTGATGGATCTGAAATGGAATGGGATTTTATGTGGCACTTGAGAAAGGT  
ACCCCGGATTGTCAGTGAAAGGACTTTCCATCTCACCAGCCCCGCATTGAGGCAGATGCTAAGATGATGGTAAA  
TACAGTGTGTGGCATCGAATGCCAGAAAGAACTCCCACTCCCAGCCTTTCTGAATTGGAGGATTATCTTTCCCTA  
TGAGACTGTCTTTGAGAAATGGCACC CGAACCTTAACCAGGGTGAAAGTTCAAGATTTGGTTCTTGAGCCGACTCA  
AAATATCACCACAAAGGGAGTATCTGTTAGGAGAAAGAGACAGGTGTATGGCACCAGCAGAGGTTGAGCATCTT  
GGACAAAAGGTTCTTAACCAATTTCCCTTTGAGCAGCTGTGAAGCTTTCCACGGGCTGTAGTGGCATTCTCAT  
TTCCCTCAGCATGTTCTAACTGCTGCCACTGTGTTTCATGATGGAAAGGACTATGTCAAAGGGAGTAAAAAGCT  
AAGGGTAGGGTTGTTGAAGATGAGGAATAAAAGTGAGGCAAGAAACGTCGAGGTTCTAAGAGGAGCAGGAGAGA  
AGCTAGTGGTGGTGACCAAAGAGAGGGTACCAGAGAGCATCTGCAGGAGAGAGCGAAGGGTGGGAGAAGAAGAAA  
AAAATCTGGCCGGGGTCAGAGGATTGCCGAAGGGAGGCCTTCCTTTGAGTGGACCCGGGTCAAGAATACCCACAT  
TCCGAAGGGCTGGGCACGAGGAGGCATGGGGGACGCTACCTTGGACTATGACTATGCTCTTCTGGAGCTGAAGCG  
TGCTCACAAAAAGAAATACATGGAACCTGGAATCAGCCCAACGATCAAGAAAATGCCTGGTGGAATGATCCACTT  
CTCAGGATTTGATAACGATAGGGCTGATCAGTTGGTCTATCGGTTTTGTCAGTGTGTCCGACGAATCCAATGATCT  
CCTTTACCAATACTGCGATGCTGAGTCGGGCTCCACCGGTTCCGGGGTCTATCTGCGTCTGAAAGATCCAGACAA  
AAAGAATTGGAAGCGCAAAATCATTGCGGTCTACTCAGGGCACCAGTGGGTGGATGTCCACGGGGTTTCAGAAGGA  
CTACAACGTTGCTGTTCCGATCACTCCCCTAAAATACGCCCAGATTTGCCTCTGGATTACAGGGAACGATGCCAA  
TTGTGCTTACGGCTAACAGAGACCTGAAACAGGGCGGTGTATCATCTAAATCACAGAGAAAACCAGCTCTGCTTA  
CCGTAGTGAGATCACTTCATAGGTTATGCCTGGACTTGAACCTCTGTCAATAGCATTTCACATTTTTCAAAATCA  
GGAGATTTTCGTCCATTTAAAAAATGTATAGGTGCAGATATTGAACTAGGTGGGCACCTCAATGCCAAGTATAT  
ACTCTTCTTTACATGGTGATGAGTTTCATTTGTAGAAAAATTTTGTTCCTTCTTAAAAATTAGACACACTTTAA  
ACCTTCAAACAGGTATTATAAATAACATGTGACTCCTTAATGGACTTATTCTCAGGGTCTACTCTAAGAAGAAAT  
CTAATAGGATGCTGGTTGTGTATTAAATGTGAAATTCATAGATAAAGGTAGATGGTAAAGCAATTAGTATCAGA  
ATAGAGACAGAAAGTTACAACACAGTTGTACTACTCTGAGATGGATCCATTCAGCTCATGCCCTCAATGTTTAT  
ATTGTGTTATCTGTTGGGTCTGGGACATTTAGTTTAGTTTTTTGAAGAATTACAAATCAGAAGAAAAAGCAAGC  
ATTATAAACAAAACATAAATACTGTTTTACTGCTTTAAGAAATAACAATTACAATGTGTATTATTTAAAAATGGGA  
GAAATAGTTTGTCTATGAAATAAACCTAGTTTAGAAATAGGGAAGCTGAGACATTTTAAGATCTCAAGTTTTTA  
TTTAATAATACTCAAAATATGGACTTTTCATGTATGCATAGGGAAGACACTTCACAAATTATGAATGATCATGT  
GTTGAAAGCCACATTATTTTATGCTATACATTCTATGTATGAGGTGCTACATTTTATAGGACAAAGAATTCTGTAA  
TCTTTTTCAAGAAAGAGTCTTTTTCTCCTTGACAAAATCCAGCTTTTGTATGAGGACTATAGGTTGAATTTCTCTG  
ATTAGTAATTTTAGATATGTCCTTTCTTAAAAATGAATAAAATTTATGAATATGA

157/615

**FIGURE 156**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA57253
<subunit 1 of 1, 413 aa, 1 stop
<MW: 47070, pI: 9.92, NX(S/T): 3
MENMLLWLIFFTPGWTLIDGSEMEWDFMWHLRKVPRIVSERTFHLTSPAFEADAKMMVNTVCGIECQKELPTPSL
SELEDYLSYETVFENGTRTLTRVKVQDLVLEPTQNITTKGVSVRRKRQVYGTDSRFSILDKRELTNFPFSTAVKL
STGCSGILISPQHVLTAACHVDGKDYVKGSKKLRVGLLKMRNKSGGKKRRGSKRSRREASGGDQREGTREHLQE
RAKGRRRRKKSGRGQRIAEGRPSFQWTRVKNTHIPKGWARGGMDATLDYDYALLELKRAHKKKYMELGISPTIK
KMPGGMIHFSGFDNDRADQLVYRFCSVSDSNLLYQYCDAESGSTSGSVYLRLLKDPDKKNWKKRIIAVYSGHQW
VDVHGVQKDYNVAVRITPLKYAQICLWIHGNDANCAYG
```

**Important features:****Signal peptide:**

amino acids 1-16

**N-glycosylation sites.**

amino acids 90-93, 110-113 and 193-196

**Glycosaminoglycan attachment site.**

amino acids 236-239

**Serine proteases, trypsin family, histidine active site.**

amino acids 165-170



158/615

**FIGURE 157**

GGGACCCATGCGGCCGTGACCCCCGGCTCCCTAGAGGCCAGCGCAGCCGAGCGGACAAAGGAGCATGTCCGCG  
CCGGGGAAGGCCGTCTCCGGCCGCCATAAGGCTCCGGTCCGCGCTGGGCCCGCGCCGCGCTCCTGCCGCCCGG  
GGCTCCGGGGCGGCCCGCTAGGCCAGTGCGCCGCCGTGCCCCCGCAGGCCCGGCCCGCAGC**ATGG**AGCCACCC  
GGACGCCGGCGGGGCCGCGCGCAGCCGCCGTGTTGCTGCCGCTCTCGCTGTTAGCGCTGCTCGCGCTGCTGGGA  
GGCGGCGGCGGCGGCGGCCGCGCGCTGCCGCCGGCTGCAAGCACGATGGGCGGCCCGAGGGGCTGGCAGG  
GCGGCGGGCGGCCCGGAGGGCAAGGTGGTGTGCAGCAGCCTGGAACTCGCGCAGGTCTGCCCCAGATACTCTG  
CCCAACCGCACGGTCACCCTGATTCTGAGTAACAATAAGATATCCGAGCTGAAGAATGGCTCATTTTCTGGGTTA  
AGTCTCCTTGAAAGATTGACCTCCGAAACAATCTTATTAGTAGTATAGATCCAGGTGCCTTCTGGGGACTGTCA  
TCTCTAAAAAGATTGGATCTGACAAACAATCGAATAGGATGTCTGAATGCAGACATATTTTCGAGGACTCACCAAT  
CTGGTTCGGCTAAACCTTTTCGGGAATTTGTTTTCTTCATTATCTCAAGGAACCTTTTGATTATCTTGCGTCATTA  
CGGTCTTTGGAATTCAGACTGAGTATCTTTGTGTGACTGTAACATACTGTGGATGCATCGCTGGGTAAAGGAG  
AAGAACATCACGGTACGGGATACCAGGTGTGTTTATCCTAAGTCACTGCAGGCCAACCACTCACAGGCGTGAAG  
CAGGAGCTGTTGACATGCGACCCTCCGCTTGAATTGCCGTCTTTTACATGACTCCATCTCATCGCCAAGTTGTG  
TTTGAAGGAGACAGCCTTCCTTTCCAGTGCATGGCTTCATATATTGATCAGGACATGCAAGTGTTGTGGTATCAG  
GATGGGAGATAGTTGAAACCGATGAATCGCAAGGTATTTTTGTTGAAAAGAACATGATTACAACCTGCTCCTTG  
ATTGCAAGTGCCCTAACCATTTCTAATATTCAGGCTGGATCTACTGGAAATTGGGGCTGTCTGTCAGACCAAAA  
CGTGGGAATAATACGAGGACTGTGGATATTGTGGTATTAGAGAGTTCTGCACAGTACTGCTCCAGAGAGGGTG  
GTAACAACAAGGTGACTTCAGATGGCCCAAGACATTGGCAGGCATTACTGCATATCTGCAGTGTACGCGGAAC  
ACCCATGGCAGTGGGATATATCCCGGAAACCCACAGGATGAGAGAAAAGCTTGGCGCAGATGTGATAGAGGTGGC  
TTTTGGGCAGATGATGATTATTCTCGCTGTCTAGTATGCAAATGATGTCACTAGAGTTCTTTATATGTTTAATCAG  
ATGCCCCCTCAATCTTACCAATGCCGTGGCAACAGCTCGACAGTTACTGGCTTACACTGTGGAAGCAGCCAACCTTT  
TCTGACAAAATGGATGTTATATTTGTGGCAGAAATGATTGAAAATTTGGAAGATTTACCAAGGAGGAAAAATCA  
AAAGAGCTAGGTGACGTGATGGTTGACATTGCAAGTAACATCATGTTGGCTGATGAACGTGTCCTGTGGCTGGCG  
CAGAGGGAAGCTAAAGCCTGCAGTAGGATTGTGCAGTGTCTTCAGCGCATTGCTACCTACCGGCTAGCCGGTGGA  
GCTCACGTTTATTCAACATATTCACCCAATATTGCTCTGGAAGCTTATGTCATCAAGTCTACTGGCTTCACGGGG  
ATGACCTGTACCGTGTTCAGAAAGTGGCAGCCTCTGATCGTACAGGACTTTCGGATTATGGGAGGCGGGATCCA  
GAGGGAACCTGGATAAGCAGCTGAGCTTTAAGTGCAATGTTTCAAATACATTTTCGAGTCTGGCACTAAAGGTA  
TGTTACATTCTGCAATCATTTAAGACTATTTACAGT**TAA**ATTAGAATGCTCCAAATGTTCTGCTTCGCAAAATAA  
CCTTATTAAAGATTTTTTTTTTGCAGGAAGATAGGTATTATTGCTTTTGCTACTGTTTTAAAGAAAACCTAACCA  
GAAGAAGTGCATTACGACTTTCAAGGGCCCTAGGCATTTTTGCTTTGATTCCCTTTCTTCACATAAAAATATCA  
GAAATTACATTTTATACTGCAGTGGTATAAATGCAAATATACTATTGTTACATGTAAAAAATTTTATTTGACT  
TAAAGTTTTATTTATTTGTTTTTTGCTCCTGATTTTAAGACAATAAGATGTTTTTCATGGGCCCTTAAAGTATC  
ATGAGCCTTTGGCACTGCGCCTGCCAAGCCTAGTGGAGAAGTCAACCTGAGACCAGGTGTTTAAATCAAGCAAGC  
TGTATATCAAAATTTTGGCAGAAAACACAAATATGTCATATATCTTTTTTAAAAAAGTATTTTATTGAAGCA  
AGCAAAATGAAAGCATTTTTTACTGATTTTTTAAATTTGGTGCTTTAGATATATTGACTACACTGTATTGAAGCAA  
ATAGAGGAGGCACAACCTCCAGCACCCCTAATGGAACCACATTTTTTTCACTTAGCTTTCTGTGGGCATGTGTAATT  
GTATTCTCTCGGTTTTTAAATCTCACAGTACTTTATTTCTGTCTTGTCCCTCAATAATATCACAAACAATATTC  
AGTCATTTTAAATGGCTGCATAATAACTGATCCAACAGGTGTTAGGTGTTCTGGTTTTAGTGTGAGCACTCAATAAA  
TATTGAATGAATGAACGAAAAAAAAAAAAAAAAA

159/615

**FIGURE 158**

MEPPGRRRRGRAQPPLLLPLSLALLALLGGGGGGGAAALPAGCKHDGRPRGAGRAAGAAEGKVVCSSLELAQVLP  
PDTLPNRTVTLLLSNNKISELKNGSFSGLSLLERLDLRNNLISSIDPGAFWGLSSLKRLDLTNNRIGCLNADIFR  
GLTNLVRNLNSGNLFSSLSQGTFDYLASLRSLFQTEYLLCDCNIIWMHRWVKEKNITVRDTRCVYPKSLQAQPV  
TGVKQELLTCDPPIELPSFYMTPSHRQVVFEGDSLFPQCMASYIDQMQVLWYQDGRIVETDESQGIFVEKNMIH  
NCSLIASALTISNIQAGSTGNWGCHVQTKRGNNTRTVDIVVLESSAQYCPPERVVNNKGDFRWPRTLGITAYLQ  
CTRNTHGSGIYPGNPQDERKAWRRCDRGGFWADDDYSRCQYANDVTRVLYMFNQMPNLNTNAVATARQLLAYTVE  
AANFSDKMDVIFVAEMIEKFGRFTKEEKSKELGDMVDIASNIMLADERVLWLAQREAKACSRIVQCLQRIATYR  
LAGGAHVYSTYSPNIALEAYVIKSTGFTGMTCTVFQKVAASDRTGLSDYGRRDPEGNLDKQLSFKCNVSNTFSSL  
ALKVCYILQSFKTIYS

**Signal peptide:**

amino acids 1-33

**Transmembrane domain:**

amino acids 13-40 (type II)

**N-glycosylation site.**amino acids 81-85, 98-102, 159-163, 206-210, 301-305, 332-336, 433-437,  
453-457, 592-596**N-myristoylation site.**amino acids 29-35, 30-36, 31-37, 32-38, 33-39, 34-40, 51-57, 57-63, 99-105,  
123-129, 142-148, 162-168, 317-323, 320-326, 384-390, 403-409, 554-560

160/615

**FIGURE 159**

GGGGAATCTGCAGTAGGCTCTGCCGGCG**ATG**GAGTGGTGGGCTAGCTCGCCGCTTCGGCTCTGGCTGCTGTTGTTTCTCCTGCCCTCAGCGCAGGGCCGCCAGAAGGAGTCAGGTTCAAAATGGAAAGTATTTATTGACCAAATTAACAGGTCTTTTGGAGAATTACGAACCATGTTCAAGTCAAAACTGCAGCTGCTACCATGGTGTATAGAAGAGGATCTAACTCCTTTCCGAGGAGGCATCTCCAGGAAGATGATGGCAGAGGTAGTCAGACGGAAGCTAGGGACCCACTATCAGATCCTAAGAACAGACTGTACCGGGAAAATGACTGCATGTTCCCTCAAGGTGTAGTGGTGTGAGCACTTTATTTTGGAAGTGATCGGGCGTCTCCCTGACATGGAGATGGTGATCAATGTACGAGATTATCCTCAGGTTCCCTAAATGGATGGAGCTGCCATCCCAGTCTTCTCCTTCAGTAAGACATCAGAGTACCATGATATCATGTATCCTGCTTGGACATTTGGGAAGGGGGACCTGCTGTTTGGCCAAATTTATCCTACAGGTCTTGGACGGTGGGACCTCTTCAGAGAAGATCTGTAAGGTGAGCAGCACAGTGGCCATGGAAAAAGAAAACTCTACAGCATATTTCCGAGGATCAAGGACAAGTCCAGAACGAGATCCTCTCATTCTTCTGTCTCGGAAAAACCCAAACTTGTTGATGCAGAATACACCAAAAACCCAGGCCTGGAAATCTATGAAAGATACCTTAGGAAAGCCAGCTGCTAAGGATGTCCATCTTGTGGATCACTGCAAATACAAGTATCTGTTTTAATTTTCGAGGCGTAGCTGCAAGTTTCCGGTTTAAACACCTCTTCCTGTGTGGCTCACTTGTTTTTCATGTTGGTGATGAGTGGCTAGAAATCTTCTATCCACAGCTGAAGCCATGGGTTCACTATATCCCAGTCAAAACAGATCTCTCCAATGTCCAAGAGCTGTTACAATTTGTAAGCAAATGATGATGTAGCTCAAGAGATTGCTGAAAGGGGAAGCCAGTTTATTAGGAACCATTTGCAGATGGATGACATCACCTGTTACTGGGAGAACCTCTTGAGTGAATACTCTAAATTCCTGTCTTATAATGTAACGAGAAGGAAAGGTTATGATCAAATTATTTCCCAAAATGTTGAAAAGTGAACCTAGTAGTCATCATAGGACCATAGTCCTCTTTGTGGCAACAGATCTCAGATATCCTACGGTGAGAAGCTTACCATAAGCTTGGCTCCTATACCTTGAATATCTGCTATCAAGCCAAATACCTGGTTTTCTTATCATGCTGCACCCAGAGCAACTCTTGAGAAAGATTTAAATGTGTCTAATACACTGATATGAAGCAGTTCAACTTTTTGGATGAATAAGGACCAGAAATCGTGAGATGTGGATTTTGAACCCAACTCTACCTTTCAATTTCTTAAGACCAATCACAGCTTGTGCCTCAGATCATCCACCTGTGTGAGTCCATCACTGTGAAATTGACTGTGTCCATGTGATGATGCCCTTTGTCCATTATTTGGAGCAGAAAATTCGTCAATTTGGAAGTAGTACAACCTATTGCTGGAATTGTGAAATTATTCAGGCGTGATCTCTGTCACTTTATTTAATGTAGGAAACCTATGGGGTTTATGAAAAATACTTGGGGATCATCTCTGAATGGTCTAAGGAAGCGGTAGCCATGCCATGCAATGATGTAGGAGTTCTCTTTGTAAAACCATAACTCTGTTACTCAGGAGGTTTCTATAATGCCACATAGAAAGAGGCCAATTGCATGAGTAATTATGCAATTGGATTTCAGGTTCCCTTTTGTGCCCTTCATGCCCTACTTCTTAATGCCTCTCTAAAGCCAAA

161/615

**FIGURE 160**

MEWWASSPLRLWLLLFLLP  
SAQGRQKESGSKWKVFIDQ  
INRSLENYEPCSSQNCSCY  
HGVIEEDLTPFRGGISRK  
MMAEVVRRKLGTHYQITKN  
RLYREND CMFPSRCSGVEH  
FILEVIGRLPDMEMVINVR  
DYPQVPKWMEPAIPVFSF  
SKTSEYHDIMYPAWTFWEG  
GPAVWPIYPTGLGRWDLFR  
EDLVRSAAQWPWKKKNST  
AYFRGSRTSPERDPLILLS  
RKNPKLVDAEYTKNQAWKS  
MKDTLGKPAAKDVHLVDH  
CKYKYLNFNFRGVAASFR  
FKHLFLCGSLVFHVGDW  
LEF  
FYPQLKPWVHYIPVKTDLS  
NVQELLQFVKANDDVAQE  
IAERGSQFIRNHLQMDDI  
TCYWENLLSEYSKFLSYN  
VT  
RRKGYDQIIPKMLKTEL

162/615

**FIGURE 161**

CCGAGCACAGGAGATTGCCCTGCGTTTAGGAGGTGGCTGCGTTGTGGGAAAAGCTATCAAGGAAGAAATTGCCAAA  
CCATGTCTTTTTTCTGTTTTTCAGAGTAGTTTACAAACAGATCTGAGTGTTTTAATTAAGCATGGAATACAGAAAA  
CAACAAAAAAGCTTAAGCTTTAATTTTCATCTGGAATTCACAGTTTTCTTAGCTCCCTGGACCCGGTTGACCTGTT  
GGCTCTTCCCGCTGGCTGCTCTATCACGTGGTGCTCTCCGACTACTCACCCGAGTGTAAGAACCCTTCGGCTCG  
CGTGCTTCTGAGCTGCTGTGGATGGCCTCGGCTCTCTGGACTGTCCTTCCGAGTAGGATGTCAGTGAATCCCTC  
AAATGGAGCCTCCTGCTGCTGTCACTCCTGAGTTTCTTTGTGATGTGGTACCTCAGCCTTCCCCACTACAATGTG  
ATAGAACGCGTGAAGTGGATGTACTTCTATGAGTATGAGCCGATTTACAGACAAGACTTTCACTTCACACTTCGA  
GAGCATTCAAAGTCTCTCATCAAAATCCATTTCTGGTCATTCTGGTGACCTCCCACCCTTCAGATGTGAAAGCC  
AGGCAGGCCATTAGAGTTACTTGGGGTGAAAAAAGTCTTGGTGGGGATATGAGGTTCTTACATTTTTCTTATTA  
GGCCAAGAGGCTGAAAAGGAAGACAAAATGTTGGCATTGTCCTTAGAGGATGAACACCTTCTTTATGGTGACATA  
ATCCGACAAGATTTTTTAGACACATATAATAACCTGACCTTGAAAACCATTATGGCATTACAGGTGGGTAAGTGA  
TTTTGCCCAATGCCAAGTACGTAATGAAGACAGACACTGATGTTTTTCATCAATACTGGCAATTTAGTGAAGTAT  
CTTTTAAACCTAAACCACTCAGAGAAGTTTTTTCACAGGTTATCCTCTAATTGATAATTATTCCTATAGAGGATTT  
TACCAAAAAACCCATATTTCTTACCAGGAGTATCCTTTCAAGGTGTCCCTCCATACTGCAGTGGGTTGGGTTAT  
ATAATGTCCAGAGATTTGGTGCCAAGGATCTATGAAATGATGGGTCACGTAAAACCCATCAAGTTTGAAGATGTT  
TATGTCGGGATCTGTTTGAATTTATTAAGTGAACATTCAATTCAGAGACACAAATCTTTTCTTCTATAT  
AGAATCCATTTGGATGTCTGTCAACTGAGACGTGTGATTGCAGCCCATGGCTTTTCTTCCAAGGAGATCATCACT  
TTTTGGCAGGTCATGCTAAGGAACACCACATGCCATTATTAACCTTCACATTCTACAAAAAGCCTAGAAGGACAGG  
ATACCTTGTGGAAAGTGTTAAATAAAGTAGGTACTGTGGAAAATTCATGGGGAGGTGAGTGTGCTGGCTTACACT  
GAAGTGAAGTCAAGAAAACCCAGACTGGAGACTGGAGGGTTACACTTGTGATTTATTAGTCAGGCCCTTCAA  
GATGATATGTGGAGGAATTAATATATAAGGAATTGGAGGTTTTTGCTAAAGAAATTAATAGGACCAACAATTTG  
GACATGTCATTCTGTAGACTAGAATTTCTTAAAGGGTGTTACTGAGTTATAAGCTCACTAGGCTGTAAAAACAA  
AACAAATGTAGAGTTTTATTATTGAACAATGTAGTCACTTGAAGGTTTTGTGTATATCTTATGTGGATTACCAAT  
TTAAAAATATATGTAGTTCTGTGTCAAAAAACTTCTTCACTGAAGTTATACTGAACAAAATTTTACCTGTTTTTG  
GTCATTTATAAAGTACTTCAAGATGTTGCAGTATTTACAGTTATTATTATTTAAATTAAGTTCACTTTGTGTT  
TTTAAATGTTTTGACGATTTCAATACAAGATAAAAGGATAGTGAATCATTCTTACATGCAACATTTTCCAGT  
TACTTAACTGATCAGTTTATTATTGATACATCACTCCATTAATGTAAAGTCATAGGTCAATTATTGCATATCAGTA  
ATCTCTTGGACTTTGTAAATATTTTACTGTGGTAATATAGAGAAGAATTAAAGCAAGAAAATCTGAAA

163/615

**FIGURE 162**

MASALWTVLPSRMSLRSLKWSLLLLSLLSFFVMWYLSLPHYNVIERVNWMYFYEYEPYRQDFHFTLREHSNC SH  
QNPFLVILVTSHPSDVKARQAIRVTWGEKKSWWGYEVLTFLLGQEA EKEDKMLALSLEDEHLLYGDII RQDFLD  
TYNNLTILKTIMAFRWVTEFCPN AKYVMKTD TDVFINTGNLVKYLNLNHSEKFFTGYPLIDNYSYRGFYQK THIS  
YQEYPFKVFPPYCSGLGYIMSRDLVPRIYEMMGHV KPIKFEDVYVGICLNLLKVNIHIPEDTNLFFLYRIHLDVC  
QLRRVIAAHGFSSKEIITFWQVMLRNTTCHY

164/615

**FIGURE 163**

CATTTCTGAACTAATCGTGTGAGAATTGACTTTGAAAAGCATTGCTTTTTACAGAAGTATATTAACTTTTTAGG  
AGTAATTTCTAGTTTGGATTGTAATATGAAATAATTTAAAAGGGCTTCGCTCATATATAGGAAAATCGCATATGG  
TCCTAGTATTAAATTCCTATTGCTTACTGATTTTTTTGAGTTAAGAGTTGTTATATGCTAGAATATGAGGATGTG  
AATATAAATAAGAGAAGAAAAAGAATAAAGTAGATTGAGTCTCCAATTTTATGTAAGCTTCAGAAGAAGTGGTT  
TGTTTACATGCAAGCTTATAGTTGAAATATTTTTCAGGAATTACATGAATGACAGTCTTCGAACCAATGTGTTTG  
TTCGATTTCAACCAGAGACTATAGCATGTGCTTGCATCTACCTTGCAGCTAGAGCACTTCAGATTCGGTTGCCAA  
CTCGTCCCCATTGGTTTTCTTCTTTTTGGTACTACAGAAGAGGAAATCCAGGAAATCTGCATAGAAACACTTAGGC  
TTTATACCAGAAAAAGCCAACTATGAATTACTGGAAAAAGAAGTAGAAAAAGAAAAGTAGCCTTACAAGAAG  
CCAAATTTAAAGCAAAGGGATTGAATCCGGATGGAACCTCCAGCCCTTCAACCCTGGGTGGATTTTCTCCAGCCT  
CCAAGCCATCATCACCAGAGAAGTAAAAGCTGAAGAGAAATCACCAATCTCCATTAATGTGAAGACAGTCAAAA  
AAGAACCTGAGGATAGACAACAGGCTTCCAAAAGCCCTTACAATGGTGTAAAGAAAAGACAGCAAGAGAAGTAGAA  
ATAGCAGAAGTGCAAGTCGATCGAGGTCAAGAACACGATCACGTTCTAGATCACATACTCCAAGAAGACACTATA  
ATAATAGGCGGAGTCGATCTGGAACATACAGCTCGAGATCAAGAAGCAGGTCCCGCAGTCACAGTGAAAGCCCTC  
GAAGACATCATAATCATGGTTCTCCTCACCTTAAGGCCAAGCATACCAGAGATGATTTAAAAAGTTCAAACAGAC  
ATGGTCATAAAAGGAAAAAATCTCGTTCTCGATCTCAGAGCAAGTCTCGGGATCACTCAGATGCAGCCAAGAAAC  
ACAGGCATGAAAGGGGACATCATAGGGACAGGCGTGAACGATCTCGCTCCTTTGAGAGGTCCCATAAAAGCAAGC  
ACCATGGTGGCAGTCGCTCAGGACATGGCAGGCACAGGCGCTGA~~CT~~CTTTCTCTTCTTTGAGCCTGCATCAGTTCT  
TGGTTTTGCCTATCTACAGTGTGATGTATGGACTCAATCAAAAACATTAAACGCAAACCTGATTAGGATTTGATTT  
CTTGAAACCCTCTAGGTCTCTAGAACACTGAGGACAGTTTCTTTTGAAAAGAACTATGTTAATTTTTTTGCACAT  
TAAAATGCCCTAGCAGTATCTAATTAAAAACCATGGTCAGGTTCAATTGTACTTTATTATAGTTGTGTATTGTTT  
ATTGCTATAAGAACTGGAGCGTGAATTCTGTAAAAATGTATCTTATTTTTATACAGATAAAATTGCAGACACTGT  
TCTATTTAAGTGGTTATTTGTTTAAATGATGGTGAATACTTTCTTAACACTGGTTTGTCTGCATGTGTAAAGATT  
TTTACAAGGAAATAAAATACAAATCCTTGTTTTTTCTAAAAAAGT

165/615

**FIGURE 164**

MNDSLRTNVFVRFPETIACACIYLAARALQIPLPTRPHWFLLFGTTEEEIQEICIETLRLYTRKKPNYELLEKE  
VEKRKVALQEAKLKAKGLNPDGTPALSTLGGFSPASKPSSPREVKAEKSPISINVKTVKKEPEDRQQASKSPYN  
GVRKDSKRSRNSRSASRSRSTRSRSRSHTPRRHYNNRRSRSGTYSSRSRSRSHSESPPRRHHNHGSPHLKAKH  
TRDDLKSSNRHGHKRKKSRSRSQSKSRDHSDAAKKHRHERGHHRDRRERSRSFERSHKSKHHGGSRSRSGHGRHR



166/615

**FIGURE 165**

GGTTCCTACATCCTCTCATCTGAGAATCAGAGAGCATAATCTTCTTACGGGCCCCGTGATTTATTAACGTGGCTTA  
ATCTGAAGGTTCTCAGTCAAATCTTTGTGATCTACTGATTGTGGGGGCATGGCAAGGTTTGCTTAAAGGAGCTT  
GGCTGGTTTTGGGCCCTTGTAGCTGACAGAAGGTGGCCAGGGAGAATGCAGCACACTGCTCGGAGAATGAAGGCGC  
TTCTGTTGCTGGTCTTGCCTTGGCTCAGTCCTGCTAACTACATTGACAATGTGGGCAACCTGCACTTCCTGTATT  
CAGAACTCTGTAAAGGTGCCTCCCACTACGGCCTGACCAAAGATAGGAAGAGGCGCTCACAAGATGGCTGTCCAG  
ACGGCTGTGCGAGCCTCACAGCCACGGCTCCCTCCCCAGAGGTTTCTGCAGCTGCCACCATCTCCTTAATGACAG  
ACGAGCCTGGCCTAGACAACCTTGCTACGTGTCTCGGCAGAGGACGGGCAGCCAGCAATCAGCCCAGTGGACT  
CTGGCCGGAGCAACCGAACTAGGGCACGGCCCTTTGAGAGATCCACTATTAGAAGCAGATCATTTAAAAAATAA  
ATCGAGCTTTGAGTGTTCTTTCGAAGGACAAAGAGCGGGAGTGCAAGTTGCCAACCATGCCGACCAGGGCAGGGAAA  
ATTCTGAAAACACCCTGCCCCGAAGTCTTTCCAAGGTTGTACCACCTGATTCCAGATGGTGAAATTACCAGCA  
TCAAGATCAATCGAGTAGATCCCACTGAAAGCCTCTCTATTAGGCTGGTGGGAGGTAGCGAAACCCCACTGGTCC  
ATATCATTATCCAACACATTTATCGTGATGGGGTGATCGCCAGAGACGGCCGGCTACTGCCAGGAGACATCATTC  
TAAAGGTCAACGGGATGGACATCAGCAATGTCCCTCACAACCTACGCTGTGCGTCTCCTGCGGCAGCCCTGCCAGG  
TGCTGTGGCTGACTGTGATGCGTGAACAGAAGTTCGCGCAGCAGGAACAATGGACAGGCCCCGGATGCCTACAGAC  
CCCGAGATGACAGCTTTTCATGTGATTCTCAACAAAAGTAGCCCCGAGGAGCAGCTTGAATAAAACTGGTGCGCA  
AGGTGGATGAGCCTGGGGTTTTTCATCTTCAATGTGCTGGATGGCGGTGTGGCATATCGACATGGTCAGCTTGAGG  
AGAATGACCGTGTGTTAGCCATCAATGGACATGATCTTCGATATGGCAGCCAGAAAGTGCGGCTCATCTGATTTC  
AGGCCAGTGAAAGACGTGTTACCTCGTGTGTCGCGCCAGGTTCCGCAGCGGAGCCCTGACATCTTTCAGGAAG  
CCGGCTGGAACAGCAATGGCAGCTGGTCCCCAGGGCCAGGGGAGAGGAGCAACACTCCCAAGCCCCCTCCATCCTA  
CAATTACTTGTGATGAGAAGGTGGTAAATATCCAAAAAGACCCCGGTGAATCTCTCGGCATGACCGTCGCAGGGG  
GAGCATCACATAGAGAATGGGATTTGCCTATCTATGTGTCATCAGTGTTGAGCCCGAGGAGTCATAAGCAGAGATG  
GAAGAATAAAAACAGGTGACATTTTGTGAATGTGGATGGGGTCGAAGTACAGAGGTCAGCCGGAGTGAGGCAG  
TGGCATTATTGAAAAGAACATCATCTCGATAGTACTCAAAGCTTTGGAAGTCAAAGAGTATGAGCCCCAGGAAG  
ACTGCAGCAGCCCAGCAGCCCTGGACTCCAACCACAACATGGCCCCACCCAGTGACTGGTCCCCATCCTGGGTCA  
TGTGGCTGGAATTACCACGGTGCTTGTATAACTGTAAAGATATTGTATTACGAAGAAACACAGCTGGAAGTCTGG  
GCTTCTGCATTGTAGGAGGTTATGAAGAATAACAATGGAAACAAACCTTTTTTTCATCAAATCCATTGTTGAAGGAA  
CACCAGCATAACAATGATGGAAGAATTAGATGTGGTGATATTCTTCTTGCTGTCAATGGTAGAAGTACATCAGGAA  
TGATACATGCTTGCTTGCAAGACTGCTGAAAGAAGTTAAAGGAAGAATTACTCTAACTATTGTTTCTTGCGCTG  
GCACTTTTTTATAGAAATCAATGATGGGTGAGAGGAAAACAGAAAAATCACAAATAGGCTAAGAAGTTGAAACACT  
ATATTTATCTTGTGATTTTTTATATTTAAAGAAAGAATACATTGTAAAAATGTGAGGAAAAGTATGATCATCTAA  
TGAAAGCCAGTTACACCTCAGAAAATATGATTCCAAAAAATTAACCTACTAGTTTTTTTTTTCAGTGTGGAGGAT  
TTCTCATTACTCTACAACATTGTTTATATTTTTTCTATTCAATAAAAAGCCCTAAAACAATAAATGATTGATT  
TGATATACCCCACTGAATCAAGCTGATTTAAATTTAAATTTGGTATATGCTGAAGTCTGCCAAGGGTACATTAT  
GGCCATTTTTTAATTTACAGCTAAAAATTTTTTAAATGCAATTGCTGAGAAACGTTGCTTTTCATCAAACAAGAAT  
AAATATTTTTTTCAGAAGTTAAA

167/615

**FIGURE 166**

MKALLLLVLPWLSPANYIDNVGNLHFLYSELCKGASHYGLTKDRKRRSQDGCPCDGCASLTATAPSPEVSAAATIS  
LMTDEPGLDNPAYVSSAEDGQPAISPVDSGRSNRTRARPFERSTIRSRSFKKINRALSVLRRTKSGSAVANHADQ  
GRESENTTAPFVFPRLYHLIPDGEITSIKINRVDPSESLSIRLVGGSETPLVHIIIQHIYRDGVIARDGRLLPG  
DIILKVNMGDISNVPHNYAVRLLRQPCQVLWLTVMREQKFRSRNNGQAPDAYRPRDDSFHVILNKSSPEEQLGIK  
LVRKVDEPGVFI FNVLDGGVAYRHGQLEENDRVLAINGHDLRYGSPESAHLIQASERRVHLVVSQRQRSPDI  
FQEAGWNSNGSWSPGPGERSNTPKPLHPTITCHEKVVNIQKDPGESLGMTVAGGASHREWDLPYVISVEPGGVI  
SRDGRIKTG DILLNVDGVELTEVSRSEAVALLKRTSSSIVLKALEVKEYEPQEDCSSPAALDSNHNMAPPSDWSP  
SWVMWLELPRCLYNCKDIVLRRNTAGSLGFCIVGGYEEYNGNKPFFIKSIVEGTPAYNDGRIRC DILLAVNGRS  
TSGMIHACLARLLKELKGRITLTIVSWPGTFL

168/615

**FIGURE 167**

GGGAAAGCCATTTGAAAACCCATCTATACAACTATATATTTTCATTTCTGCTGCTAGCTGCCTTGGGCCTCAC  
AATTTTCATTCTGTTTTCTGACTTTCAAGTTATATACCGTGGAATGGAGTTGATCCCAACCATAACATCGTGGAG  
GGTTTTAATTTGGTGGTAGCCCTCACCCAATTCTGGTGTGGCTTTCTTGCAGAGGATTCCACCTTCAAATCA  
TGAACCTCTGGCTGTTGATCAAAAGAGAATTTGGATTCTACTCTAAAAGTCAATATAGGACTTGGCAAAAGAAGCT  
AGCAGAAGACTCAACCTGGCCTCCCATAAACAGGACAGATTATTCAGGTGATGGCAAAAATGGATTCTACATCAA  
CGGAGGCTATGAAAGCCATGAACAGATTCCAAAAGAAAACCTCAAATGGGAGGCCAACCCACAGAACAGCATT  
CTGGGCCAGGCTGTAAATCAGAATTGTCGTCGTACATGCTCAACAGCATTGCTTTTTTCCCCAAAATTAACACATT  
GTGGAGAAGTGATGATACTCTCCCCTTACCTTTCCTCTCTCCATTCAAGCATTCAAAGTATATTTTCAATGAATT  
AAACCTTGCAGCAAGGGACCTTAGATAGGCTTATTCTGACTGTATGCTTTACCAATGAGAGAAAAAATGCATT  
CCTGTATCATCCTTTTCAATAAACTGTATTCATTTTGAAAAAAAAAAAAAAAAAAAAA

169/615

**FIGURE 168**

MELIPTITSWRVLILVVALTQFWCGFLCRGFHLQNHFWLLIKREFGFYSKSQYRTWQKLAEDSTWPPINRTDY  
SGDGKNGFYINGGYESHEQIPKRKLKLGGPTEQHFWARL

**FIGURE 169**

[illegible]

171/615

**FIGURE 170**

MELGCWTQLGLTFLQLLLISSLPREYTVINEACPGAENIMCRECCEYDQIECVCPGKREVVGYTIPCCRNEENE  
CDSCLIHPGCTIFENCKSCRNGSWGGLDDEFYVKGFYCAECRAGWYGGDCMRGQVLRAPKGQILLESYPLNAHC  
EWTIHAKPGFVIQLRFVMLSLEFDYMCQYDYVEVRDGDNRDGGQIIKRVCGNERPAPIQSIGSSSLHVLFHSDGSKN  
FDGFHAIYEEITACSSSPCFHDGTCVLDKAGSYKCACLAGYTGQRCENLLEERNCSDPGGPVNGYQKITGGPGLI  
NGRHAKIGTVVSFFCNSYVLSGNEKRTCQQNGEWSGKQPICIKACREPKISDLVRRRVLPQVQSRETPLHQLY  
SAAFSKQKLQSAPTKKPALPFGDLPMGYQHLHTQLQYECISPFYRRLGSSRRTCLRTGKWSGRAPSCIPICGKIE  
NITAPKTQGLRWPWQAAYRRTSGVHDGSLHKGAWFLVCSGALVNERTVVVAHCVTDLGKVTMIKTADLKVVLG  
KFYRDDDREKTIQSLQISAILHLPNYDPILLDADIAILKLLDKARISTRVQPICLAASRDLSTSFQESHITVAG  
WNVLADVRSFGFKNDTLRSGVSVVDSLLCEEQHEDHGIPVSVTDNMFCASWEPTAPSDICTAETGGIAAVSFPG  
RASPEPRWHLMLVSWSYDKTCSHRLSTAFTKVLPFKDWIERNMK

172/615

**FIGURE 171**

CTGTCGTCTTTGCTTCAGCCGCAGTCGCCACTGGCTGCCTGAGGTGCTCTTACAGCCTGTTCCAAGTGTGGCTTA  
ATCCGTCTCCACCACCAGATCTTTCTCCGTGGATTCTCTGCTAAGACCGCTGCC**ATG**CCAGTGACGGTAACCCG  
CACCACCATCACAACCACCACGACGTCATCTTCGGGCCTGGGGTCCCCATGATCGTGGGGTCCCCTCGGGGCCCT  
GACACAGCCCCCTGGGTCTCCTTCGCCTGCTGCAGCTGGTGTCTACCTGCGTGGCCTTCTCGCTGGTGGCTAGCGT  
GGGCGCCTGGACGGGGTCCATGGGCAACTGGTCCATGTTACCTGGTGTCTCTGCTTCTCCGTGACCCGTATCAT  
CCTCATCGTGGAGCTGTGCGGGCTCCAGGCCCGCTTCCCCCTGTCTTGGCGCAACTTCCCCATCACCTTCGCCTG  
CTATGCGGCCCTCTTCTGCCTCTCGGCCTCCATCATCTACCCACCACCTATGTCCAGTTCCTGTCCACGGCCG  
TTCGCGGGACCACGCCATCGCCGCCACCTTCTTCTCCTGCATCGCGTGTGTGGCTTACGCCACCGAAGTGGCCTG  
GACCCGGGCCCGGGCCCGGCGAGATCACTGGCTATATGGCCACCGTACCCGGGCTGCTGAAGGTGCTGGAGACCTT  
CGTTGCCTGCATCATCTTCGCGTTCATCAGCGACCCCAACCTGTACCAGCACCAGCCGGCCCTGGAGTGGTGCCT  
GGCGGTGTACGCCATCTGCTTCATCCTAGCGGCCATCGCCATCCTGCTGAACCTGGGGGAGTGCACCAACGTGCT  
ACCCATCCCCCTCCCCAGCTTCCTGTGCGGGCTGGCCTTGCTGTCTGTCTCTCTATGCCACCGCCCTTGTTCCT  
CTGGCCCCCTCTACCAGTTCGATGAGAAGTATGGCGGCCAGCCTCGGCGCTCGAGAGATGTAAGCTGCAGCCGCAG  
CCATGCCTACTACGTGTGTGCCTGGGACCGCCGACTGGCTGTGGCCATCCTGACGGCCATCAACCTACTGGCGTA  
TGTGGCTGACCTGGTGCATCTGCCCCACCTGGTTTTTGTCAAGGTCT**AA**AGACTCTCCCAAGAGGCTCCCGTTCCC  
TCTCCAACCTCTTTGTTCTTCTTGCCCGAGTTTTCTTTATGGAGTACTTCTTTCCCTCCGCCCTTTCCTCTGTTTC  
CTCTTCTGTCTCCCTCCCTCCACCTTTTTCTTTCTTCCCAATTCCCTGCACTCTAACCAGTTCCTGGATGC  
ATCTTCTTCCCTCCCTTTCTCTTGCTGTTTCTTCTCTGTGTTTGTGTTTGGTGGCCACATCCTGTTTTCACCCCTG  
AGCTGTTTCTCTTTTCTTTTCTTTTCTTTTTTTTTTTTTTTTTTAAGACGGATTCTCACTCTGTGGCCCAGGCTG  
GAGTGCAGTGGTGCATCTCAGCTCACTGCAACCCCCGCCCTCCTGGGTTCAGCGATTCTCCTCCCCCAGCCTCC  
CAAGTAGCTGGGAGGACAGGTGTGAGCTGCCGCACCCAGCCTGTTTCTCTTTTCCACTCTTCTTTTTTCTCATC  
TCTTTTCTGGGTGCTGTCGGCTTCTTATCTGCCTGTTTGAAGCACCTTCTCCTGTGTCCTTGGGAGCCCT  
GAGACTTCTTTCTCTCCTTGCTCCACCCACCTCCAAAGGTGCTGAGCTCACATCCACACCCCTTGCAGCCGTCC  
ATGCCACAGCCCCCAAGGGGCCCCATTGCCAAAGCATGCCTGCCACCCCTCGCTGTGCCTTAGTCAGTGTGTAC  
GTGTGTGTGTGTGTGTTGGGGGGTGGGGGGTGGGTAGCTGGGGATTGGGCCCTCTTCTCCAGTGGAGGAA  
GGTGTGCAGTGTACTTCCCCTTTAAATTAAAAACATATATATATATATATTTGGAGGTGAGTAATTTCCAATGG  
GCGGGAGGCATTAAGCACCGACCTGGGTCCCTAGGCCCGCCTGGCACTCAGCCTTGCCAGAGATTGGCTCCAG  
AATTTTTGCCAGGCTTACAGAACACCCACTGCCTAGAGGCCATCTTAAAGGAAGCAGGGGCTGGATGCCTTTCAT  
CCCAACTATTCTCTGTGATGAAAAAG

173/615

**FIGURE 172**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58727
<subunit 1 of 1, 322 aa, 1 stop
<MW: 35274, pI: 8.57, NX(S/T): 1
MPVTVTRTTITTTTSSSGLGSPMIVGSPRALTQPLGLLRLLQLVSTCVAFSLVASVGAWTGSMGNWSMFTWCFC
FSVTLLIILIVELCGLQARFPLSWRNFPIITFACYAALFCLSASIIYPTYVQFLSHGRSRDHAIATFFSCIACVA
YATEVAWTRARPGEITGYMATVPGLLKVLETFVACIIFAFISDPNLYQHQPALWCVAVYAICFILAAIAILLNL
GECTNVLPIPFPSFLSGLALLSVLLYATALVLWPLYQFDEKYGGQPRRSRDVSCSRSHAYYVCAWDRRLAVAILT
AINLLAYVADLVHSAHLVFKV
```

**Important features:****Transmembrane domains:**

amino acids 41-60 (type II), 66-85, 101-120, 137-153, 171-192, 205-226, 235-255 and 294-312

**N-glycosylation site.**

amino acids 66-69

**Glycosaminoglycan attachment site.**

amino acids 18-21



174/615

**FIGURE 173**

GAACGTGCCACCATGCCAGCTAATTTTTGTATTTTTAGTAGAGACGGGGTTTCACCATGTTGGCCAGGCTGGTCT  
TTGAACTCGTGACCTCATGATCCGCTCACCTCGGCCTCCCAAAGTGCTGGGATTACAGGCATGAGCCACTGACGC  
CTGGCCAGCCTATGCATTTTTTAAGAAATTATTTCTGTATTAGGTGCTGTGCTAAACATTGGGCACAGTGACCA  
AAACAGACTGAATTTCCCAAGAGCCAAAGACCAGTGAGGGAGACCAACAAGAAACAGGAAATGCAAAAGAGACCA  
TTATTACTCACTATGACTAAGGGTCACAAATGGGGTACGTTGATGGAGAGTGATTTGTTAAGAGACTACAGAGGG  
AGGACAGACTACCAAGAGGGGGGCCAGGAAAGCTCCTCTGACGAGGTGGTATTTTCAGCCCAAAGTGAAGAATGA  
GAAAGAGCTAGCCAGCCATCAGAATAGTCCAGAAGAGATGGGGAGCACTACACTCACTACACTTTGGCCTGAGAA  
AATAGCATGGGATTGGAGGAGGCTGGGGGAACACCCTTCTGCCGACCTGGGCAGGAGGCATTGAGGGCTTGAGA  
AAGGGCAATGGCAGTAGCAGTAGAAAGGACAGGGTAGGAGCAGGGACTTTGCAGGTGGAATCATTAGGTCTTATC  
AACAGATATGGGCAAGCAAAGCCAGGGGAGAATTGATGGTAATGCTGAGGTTTGGAGCCAGGCTAGATGGGACAG  
TGGTGGGTGATGCAAAGGAAAGAGGTGAGGAAGCAGGGCCAGACGTGGGGAGAAGGTGTGGGGGTTTGGTTTCCA  
TCTTGCCGAGTCTGCCGGAATGTGGATGGGAAGACCAAGAGGAGGAGCAAGGGGCAGAGGGGAAGGGAATCTTAA  
AGAAGTCTGGATGCCACACTCTTCTCCTTCCCTCCTCCTCCTCCTCCTCCTCAGAGGTCTCACTCGTGGTCTTCTCAT  
TTCTGCCCCTGCCTCCATCTCCTCTGGGTGCTGGGAAAGTGGAGGATTAGCTGAAGTTTTGCTTCTCGGGGCCCTG  
TCTGAATCTCCATTGCTTTCTGGGAGGACATAATTACCTGTCTTAGCTTCTTATCATCTTACATTTCCCTGTAG  
CCACTGGGACATATGTGTGTTCCTTCCCTAGCTCCTGTCTCCTCCTCATGCTTTGCTGGGTATGGGCATGTTAG  
GGGGAAGGTCAATTGCTGTGAGAGGGGCACTGACTTTCTAATGGTGTTACCCAAGGTGAATGTTGGAGACACAGTC  
GCGATGCTGCCAAGTCCCGGCGAGCCCTAACTATCCAGGAGATCGCTGCGCTGGCCAGGTCTCCCTGCATGGT  
ATGCAGCCCCCTCCCATGTTTCTGGCCACTTTGTCCTTTCTCCTCCCCTTTCACATCCCTTTGGAAGTGTTCCT  
GTGAGTACATGCTGGGGTCTCCCTTTCTTCCCTTGCTCAGGTGAATCTCAGCCCCCTTCTCCCAACCAAGGTTT  
ACATGGATCTTAACACTGCCACCCCTTCCACCTCCCTGCACCTGTGCTCCCTGGCCTGGTCTTTTACCAGGCTTC  
TCCACCCCTCCCCTATCTCCAGGTATTTCCAGGTGGTGAAGGACCAAGTGAACCAAGCCTACCGCATGGCCAGG  
GCCGAGTGGCTCACCTCATTGAGTGGAAAGGGCTGGAGCAAGCCGAGTGACTCACCTGCTGCCCTGGAATCAGCCT  
TTTCTCCTATTTCAGACCTCAGCGAGGGCGAACAAGAGGCTCGCTTTGCAGCAGGAGTGGCTGAGCAGTTTGCCA  
TCGCGGAAGCCAAGCTCCGAGCATGGTCTTCGGTGGATGGCGAGGACTCCACTGATGACTCCTATGATGAGGACT  
TTGCTGGGGGAATGGACACAGACATGGCTGGGCAGCTGCCCTGGGGCCGACCTCCAGGACCTGTTACACGGGCC  
ACCGGTTCTCCCGGCTGTGCGCCAGGGCTCCGTGGAGCCTGAGAGCGACTGCTCACAGACCGTGTCCCCAGACA  
CCCTGTGCTCTAGTCTGTGCAGCCTGGAGATGGGTGTTGGGGTCCCCGGCCGGCTGGCCTCCAGCTGCTGG  
GCGATGAGCTGCTTCTCGCCAACTGCCCCCCAGCCGGGAAAGTGCCTTCCGCAGCCTGGGCCCACCTGGAGGCC  
AGGACTCACTCTACAACCTCGCCCCCTCACAGAGTCTGCTTTTCCCCCGGAGGAGGAGCCAGCCCCCTGCAAGG  
ACTGCCAGCCACTCTGCCACCACTAACGGGCAGCTGGGAACGGCAGCGCAAGCCTCTGACCTGGCCTCTTCTG  
GGGTGGTGTCTTAGATGAGGATGAGGCAGAGCCAGAGGAACAGTGAACCCACATCATGCCTGGCAGTGGCATGCA  
TCCCCCGGCTGCTGCCAGGGGCAGAGCCTCTGTGCCAAGTGTGGGCTCAAGGCTCCCAGCAGAGCTCCACAGCC  
TAGAGGGCTCCTGGGAGCGCTCGCTTCTCCGTGTGTGTTTTGCATGAAAGTGTGAGAGAGGAGGAGGGGCTG  
GGCTGGGGGCGCATGTCTGCCCCCACTCCCGGGGCTTGCCGGGGGTTGCCGGGGGCTCTGGGGCATGGCTACA  
GCTGTGGCAGACAGTGATGTTTCTTAAATGCCACACACACATTTCTCCTCGGATAATGTGAACCACTA  
AGGGGGTTGTGACTGGGCTGTGTGAGGGTGGGGTGGGAGGGGGCCAGCAACCCCCCACCCTCCCCATGCCTCTC  
TCTTCTCTGCTTTTCTTCTCACTTCCGAGTCCATGTGCACTGCTTGATAGAATCACCCCCACCTGGAGGGGCTGG  
CTCCTGCCCTCCCGGAGCCTATGGGTTGAGCCGTCCCTCAAGGGCCCCCTGCCAGCTGGGCTCGTGCTGTGCTTC  
ATTACCTCTCCATCGTCTCTAAATCTTCTTCTTTTCTTAAAGACAGAAGGTTTTTGGTCTGTTTTTTCAGTC  
GGATCTTCTCTTCTCTGGGAGGCTTTGGAATGATGAAAGCATGTACCCCTCCACCCTTTTCTGGCCCCCTAATGG  
GGCCTGGGGCCCTTTCCCAACCCCTCCTAGGATGTGCGGGCAGTGCTGCTGGCGCCTCACAGCCAGCCGGGCTGCC  
ATTCACGCAGAGCTCTCTGAGCGGGAGGTGGAAGAAAGGATGGCTCTGGTTGCCACAGAGCTGGGACTCATGTT  
CTTCTAGAGAGGGCCACAAGAGGGCCACAGGGGTGGCCGGGAGTTGTGAGCTGATGCCTGCTGAGAGGCAGGAAT  
TGTGCCAGTGAGTGACAGTCATGAGGGAGTGTCTCTTGGGGAGGAAAGAGGTAGAGCCTTTCTGTCTGAAT  
GAAAGGCCAAGGCTACAGTACAGGGCCCCCGCCAGCCAGGGTGTAAATGCCACGTAAGTGGAGGCCCTCTGGCAG  
ATCTGCATTTCAAGGTCACTGGACTGTACGTTTTTATGTTGTGGGAAGGGTGGGTGGTATTAAGGGC  
CTTGTAGGCTTTGGCAGGTAAGAGGGCCCAAGGTAAGAACGAGAGCAACGGGCACAAGCATTCTATATATAAGT  
GGCTCATTAGGTGTTTATTTTGTCTATTTAAGAATTTGTTTTATTAAATTAATATAAAAATCTTGTAAATCTC  
TAAAA

175/615

**FIGURE 174**

MFLATLSFLLPFAHPFGTVSCEYMLGSPLSSLAQVNLSPFSPKPVHMDPNYCHPSTSLHLCSLAWSFTRLLHPPL  
SPGISQVVKDHVTKPTAMAQGRVAHLIEWKGWSKPSDSPAALESASFSSYSDLSEGEQEARFAAGVAEQFAIAEAK  
LRAWSSVDGEDSTDDSYDEDFAGGMDTDMAGQLPLGPHLQDLFTGHRFSRPVRQGSVEPESDCSQTVPD TLCSS  
LCSLEDGLLGSPARLASQLLGDELLAKLPPSRESAFRSLGPLEAQDSLNSPLTESCLSPAEEEPAPCKDCQPL  
CPPLTGSWERQRQASDLASSGVVSLDEDEAEPEEQ

**Signal peptide:**

amino acids 1-15

**Casein kinase II phosphorylation site.**amino acids 123-127, 128-132, 155-159, 162-166, 166-170, 228-232, 285-289,  
324-328**Tyrosine kinase phosphorylation site.**

amino acids 44-52

**N-myristoylation site.**

amino acids 17-23, 26-32, 173-179

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 11-22

176/615

**FIGURE 175**

GGTTCCTGGGCGCTCTGTTACACAAGCAAGATACAGCCAGCCCCACCTAATTTTGTTCCTGGCACCCCTCCTGC  
TCAGTGCGACATTGTACACTTAACCCATCTGTTTTCTCTAATGCACGACAGATTCTTTTCAGACAGGACAACCTG  
TGATATTTTCAGTTCCTGATTGTAAATACCTCCTAAGCCTGAAGCTTCTGTTACTAGCCATTGTGAGCTTCAGTTT  
CTTCATCTGCAAAATGGGCATAATACAATCTATTCTTGCCACATCAAGGGATTGTTATTCCTTTAAAAAAAACC  
AATACCAAAGAAGCCTACAATGTTGGCCTTAGCCAAAATTCTGTTGATTTCACGTTGTTTTATTCACTTCTATC  
GGGGAGCCATGGAAAAGAAAATCAAGACATAAACACAACACAGAACATTGCAGAAGTTTTTAAACAATGGAAAA  
TAAACCTATTTCTTTGGAAAGTGAAGCAAACCTTAAACTCAGATAAAGAAAATATAACCACCTCAAATCTCAAGGC  
GAGTCATTCCCCTCCTTTGAATCTACCCAACAACAGCCACGGAATAACAGATTTCTCCAGTAACTCATCAGCAGA  
GCATTCTTTGGGCAGTCTAAAACCCACATCTACCATTTCCACAAGCCCTCCCTTGATCCATAGCTTTGTTTCTAA  
AGTGCCTTGGAAATGCACCTATAGCAGATGAAGATCTTTTGCCCATCTCAGCACATCCCAATGCTACACCTGCTCT  
GTCTTCAGAAAACCTTCACTTGGTCTTTGGTCAATGACACCGTGAAAACCTCCTGATAACAGTTCATTACAGTTAG  
CATCCTCTCTTCAGAACCAACTTCTCCATCTGTGACCCCTTGATAGTGGAACCAAGTGGATGGCTTACCACAAA  
CAGTGATAGCTTCACTGGGTTTACCCCTTATCAAGAAAAAACAACCTCTACAGCCTACCTTAAAATTACCAATAA  
TTCAAAACTCTTTCAAATACGTGAGATCCCCAAAAGAAAATAGAAATACAGGAATAGTATTCGGGGCCATTTT  
AGGTGCTATTCTGGGTGTCTCATTGCTTACTCTTGTGGGCTACTTGTGTGTGGAAAAAGGAAAACGGATTCAAT  
TTCCCATCGGCGACTTTATGACGACAGAAATGAACAGTTCTGCGATTAGACAATGCACCGGAACCTTATGATGT  
GAGTTTTGGGAATTCTAGCTACTACAATCCAACCTTTGAATGATTAGCCATGCCAGAAAGTGAAGAAAATGCACG  
TGATGGCATTTCCTATGGATGACATACCTCCACTTCGTACTTCTGTATAGAACTAACAGCAAAAAGGCGTTAAACA  
GCAAGTGTCTATCTACATCCTAGCCTTTTGACAAATTCATCTTTCAAAGGTTACACAAAATTACTGTCACGTGGA  
TTTTGTCAAGGAGAATCATAAAGCAGGAGACCAGTAGCAGAAATGTAGACAGGATGTATCATCAAAGGTTTTT  
TTTCTTACAATTTTTGGCCATCCTGAGGCATTTACTAAGTAGCCTTAATTTGTATTTTAGTAGTATTTTCTTAGT  
AGAAAAATATTTGTGGAATCAGATAAACTAAAAGATTTCAACATTACAGCCCTGCCTCATAACTAAATAATAAAA  
ATTATTCCACCAAAAAATTTCTAAAACAATGAAGATGACTCTTTACTGCTCTGCCTGAAGCCCTAGTACCATAATT  
CAAGATTGCATTTTCTTAAATGAAAATTGAAAGGGTGCTTTTTTAAAGAAAATTTGACTTAAAGCTAAAAAGAGGA  
CATAGCCCAGAGTTTCTGTTATTGGGAAATTGAGGCAATAGAAATGACAGACCTGTATTCTAGTACGTTATAATT  
TTCTAGATCAGCACACATGATCAGCCCACTGAGTTATGAAGCTGACAATGACTGCATTCAACGGGGCCATGGC  
AGGAAAGCTGACCCTACCCAGGAAAGTAATAGCTTCTTTAAAAGCTTCAAAGGTTTTGGGAATTTTAACTTGTC  
TTAATATATCTTAGGCTTCAATTATTTGGGTGCCTTAAAACTCAATGAGAATCATGGT

177/615

**FIGURE 176**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58732
><subunit 1 of 1, 334 aa, 1 stop
><MW: 36294, pI: 4.98, NX(S/T): 13
MLALAKILLISTLFYSLLSGSHGKENQDINTTQNI AEVFKTMENKPISLESEANLNSDKENITTSNLKASHSPPL
NLPNNSHGITDFSSNSSAEHSLGSLKPTSTISTSPPLIHSFVSKVPWNAPIADEDLLPISAHPNATPALSSSENF
WSLVNDTVKTPDNSSITVSILSSEPTSPSVTPLIVEPSGWLTNNSDSFTGFTPYQEKTTLQPTLKFTNNSKLF
TSDPQKENRNTGIVFGAILGAILGVSLTLVGYLLCGKRKTD SFSHRRLYDDRNEPVLRLDNAPEPYDV SFGNSS
YYNPTLNSAMPESEENARDGIPMDDIPPLRTSV
```

**Signal peptide:**

amino acids 1-23

**Transmembrane domain:**

amino acids 235-262

**N-glycosylation site.**amino acids 30-34, 61-65, 79-83, 90-94, 148-152, 155-159, 163-167, 218-222,  
225-229, 298-302, 307-311

178/615

**FIGURE 177**

ACCAGGCATTGTATCTTCAGTTGTCATCAAGTTCGCAATCAGATTGGAAAAGCTCAACTTGAAGCTTTCTTGCCCT  
GCAGTGAAGCAGAGAGATAGATATTATTACGTAATAAAAAACATGGGGCTTCAACCTGACTTTCCACCTTTCCCTA  
CAAATTCGATTACTGTTGCTGTTGACTTTGTGCCCTGACAGTGGTTGGGTGGGCCACCAGTAACTACTTCGTGGG  
TGCCATTCAAGAGATTCCCTAAAGCAAAGGAGTTCATGGCTAATTTCCATAAGACCCTCATTTTGGGGAAGGGAAA  
AACTCTGACTAATGAAGCATCCACGAAGAAGGTAGAAGTTGACAACTGCCTTCTGTGTCTCCTTACCTCAGAGG  
CCAGAGCAAGCTCATTTTCAAACCAGATCTCACTTTGGAAGAGGTACAGGCAGAAAATCCCAAAGTGTCCAGAGG  
CCGGTATCGCCCTCAGGAATGTAAAGCTTTACAGAGGGTCCCATCCTCGTTCCCCACCGGAACAGAGAGAAAACA  
CCTGATGTACCTGCTGGAACATCTGCATCCCTTCTGCAGAGGCAGCAGCTGGATTATGGCATCTACGTCATCCA  
CCAGGCTGAAGGTAAAAAGTTTAAATCGAGCCAACTCTTGAATGTGGGCTATCTAGAAGCCCTCAAGGAAGAAA  
TTGGGACTGCTTTATATTCCACGATGTGGACCTGGTACCCGAGAATGACTTTAACCTTTACAAGTGTGAGGAGCA  
TCCCAAGCATCTGGTGGTTGGCAGGAACAGCACTGGGTACAGGTTACGTTACAGTGGATATTTTGGGGGTGTTAC  
TGCCCTAAGCAGAGAGCAGTTTTTCAAGGTGAATGGATTCTCTAACAACACTACTGGGATGGGGAGGCGAAGACGA  
TGACCTCAGACTCAGGGTTGAGCTCCAAAGAAATGAAAATTTCCCGGCCCTGCCTGAAGTGGGTAAATATACAAT  
GGTCTTCCACACTAGAGACAAAGGCAATGAGGTGAACGCAGAACGGATGAAGCTCTTACACCAAGTGTACAGAGT  
CTGGAGAACAGATGGGTTGAGTAGTTGTTCTTATAAATTAGTATCTGTGGAACACAATCCTTTATATATCAACAT  
CACAGTGGATTTCTGGTTGGTGCATGACCCCTGGATCTTTTGGTGATGTTTGGGAAGAACTGATTCTTTGTTTGCA  
ATAATTTTGGCCTAGAGACTTCAAATAGTAGCACACATTAAAGAACCTGTTACAGCTCATTGTTGAGCTGAATTTT  
TCCTTTTTGTATTTTCTTAGCAGAGCTCCTGGTGATGTAGAGTATAAAACAGTTGTAACAAGACAGCTTTCTTAG  
TCATTTTGATCATGAGGGTTAAATATTGTAATATGGATACTTGAAGGACTTTATATAAAAGGATGACTCAAAGGA  
TAAATGAACGCTATTTGAGGACTCTGGTTGAAGGAGATTTATTTAAATTTGAAGTAATATATTATGGGATAAAA  
GGCCACAGGAAATAAGACTGCTGAATGTCTGAGAGAACCAGAGTTGTTCTCGTCCAAGGTAGAAAGGTACGAAGA  
TACAATACTGTTATTCATTTATCCTGTACAATCATCTGTGAAGTGGTGGTGTGAGGTGAGAAGGCGTCCACAAAA  
GAGGGGAGAAAAGGCGACGAATCAGGACACAGTGAAGTGGGAATGAAGAGGTAGCAGGAGGGTGGAGTGTCCGC  
TGCAAAGGCAGCAGTAGCTGAGCTGGTTGCAGGTGCTGATAGCCTTCAGGGGAGGACCTGCCAGGTATGCCTTC  
CAGTGATGCCACCAGAGAATACATTCTCTATTAGTTTTTAAAGAGTTTTTGTAAATGATTTTGTACAAGTAGG  
ATATGAATTAGCAGTTTACAAGTTTACATATTAATAATAAATATGTCTATCAAATACCTCTGTAGTAAAT  
GTGAAAAAGCAAAA

179/615

**FIGURE 178**

MGFNLT FHLSYKFRLLLLLLTLCITVVGWATSNYFVGAIQEIPKAKEFMANFHKTLILGKGKTLTNEASTKKVELD  
NCPSVSPYLRGQSKLIFKPDLTLEEVQAENPKVSRGRYRPQECKALQRVAILVPHRNREKHLMYLLEHLHPFLQR  
QQLDYGIYVIHQAEKGKFNRAKLLNVGYLEALKEENWDCFI FHDVDLVPENDFNLYKCEEHPKHLVVGRNSTGYR  
LRYSGYFGGVTA LSREQFFKVNGFSNNYWGWGGEDDLRLRLVELQRMKISRPLPEVGKYTMVFHTRDKGNEVNAE  
RMKLLHQVSRVWRTDGLSSCSYKLVSV EHNPLYINITVDFWFGA

**Important features:****Signal peptide:**

amino acids 1-27

**N-glycosylation sites.**

amino acids 4-8, 220-224, 335-339

**Xylose isomerase proteins.**

amino acids 191-202

180/615

**FIGURE 179**

CGTGGGCCGGGGTTCGCGCAGCGGGCTGTGGGCGCGCCCGGAGGAGCGACCGCCGCAGTTCTCGAGCTCCAGCTGC  
ATTCCCTCCGCGTCCGCCCCACGCTTCTCCCGCTCCGGGCCCGCAATGCCCCAGGCAGTGTGGTTCGCGCCTCGG  
CCGCATCCTCTGGCTTGCCCTGCCCTGCCCTGGGCCCGGCAGGGGTGGCCGAGGCCCTGTATGAACTCAATCT  
CACCACCGATAGCCCTGCCACCACGGGAGCGGTGGTGACCATCTCGGCCAGCCTGGTGGCCAAGGACAACGGCAG  
CCTGGCCCTGCCCGCTGACGCCACCTCTACCGCTTCCACTGGATCCACACCCCGCTGGTGGCTTACTGGCAAGAT  
GGAGAAGGGTCTCAGTCCACCATCCGTGTGGTGGGCCAGTGGCCGGGGAATTCGGGTCTCTGTCTGGGTAC  
TGCCGCTGACTGCTGGATGTGCCAGCCTGTGGCCAGGGGCTTTGTGGTCCTCCCCATCACAGAGTTCCCTCGTGGG  
GGACCTTGTTGTACCCAGAACACTTCCCTACCCTGGCCAGCTCCTATCTCACTAAGACCGTCCCTGAAAGTCTC  
CTTCCTCCTCCAGACCCGAGCAACTTCCCTCAAGACCGCCTTGTTCCTCTACAGCTGGGACTTCGGGGACGGGAC  
CCAGATGGTGACTGAAGACTCCGTGGTCTATTATAACTATTCCATCATCGGGACCTTCACCGTGAAGCTCAAAGT  
GGTGGCGGAGTGGGAAGAGGTGGAGCCGGATGCCACGAGGGCTGTGAAGCAGAAGACCGGGGACTTCTCCGCCTC  
GCTGAAGCTGCAGGAAACCTTCGAGGCATCCAAGTGTGGGGCCACCCTAATTCAGACCTTCCAAAAGATGAC  
CGTGACCTTGAACCTCCTGGGGAGCCCTCCTCTGACTGTGTGCTGGCGTCTCAAGCCTGAGTGCCTCCCGCTGGA  
GGAAGGGGAGTGCCACCCCTGTGTCCGTGGCCAGCACAGCGTACAACCTGACCCACACCTTCAGGGACCCCTGGGGA  
CTACTGCTTCAGCATCCGGGCCGAGAATATCATCAGCAAGACACATCAGTACCACAAGATCCAGGTGTGGCCCTC  
CAGAATCCAGCCGGCTGTCTTTGCTTTCCCATGTGCTACACTTATCACTGTGATGTTGGCCTTCATCATGTACAT  
GACCTGCGGAATGCCACTCAGCAAAAGGACATGGTGGAGAACCCGGAGCCACCCTCTGGGGTCAGGTGCTGCTG  
CCAGATGTGCTGTGGGCCTTTCTTGCTGGAGACTCCATCTGAGTACCTGGAAATTGTTGCTGAGAACACGGGCT  
GCTCCCGCCCTCTATAAGTCTGTCAAACTTACACCGTGTGAGCACTCCCCCTCCCCACCCCATCTCAGTGTTA  
ACTGACTGCTGACTTGGAGTTTCCAGCAGGGTGGTGTGCACCACTGACCAGGAGGGGTTCAATTTGCGTGGGGCTG  
TTGGCCTGGATCATCCATCCATCTGTACAGTTCAGCCACTGCCACAAGCCCCCTCCCTCTCTGTACCCCTGACCC  
CAGCCATTCACCCATCTGTACAGTCCAGCCACTGACATAAGCCCCACTCGGTTACCACCCCTTGACCCCTACC  
TTTGAAGAGGCTTCGTGCAGGACTTTGATGCTTGGGGTGTTCGCTGTGACTCCTAGGTGGGCCTGGCTGCCAC  
TGCCCATTCCTCTCATATTGGCACATCTGCTGTCCATTGGGGGTTCCTCAGTTTCCTCCCCCAGACAGCCCTACCT  
GTGCCAGAGAGCTAGAAAGAAGGTCATAAAGGGTTAAAAATCCATAACTAAAGGTTGTACACATAGATGGGCACA  
CTCACAGAGAGAAGTGTGCATGTACACACACACACACACACACACACACAGAAATATAAACACATG  
CGTCACATGGGCATTTAGATGATCAGCTCTGTATCTGGTTAAGTCGGTTGCTGGGATGCACCCCTGCACTAGAGC  
TGAAAGGAAATTTGACCTCCAAGCAGCCCTGACAGGTTCTGGGCCCGGGCCCTCCCTTTGTGCTTTGTCTCTGCA  
GTTCTTGCGCCCTTTATAAGGCCATCCTAGTCCCTGCTGGCTGGCAGGGGCTGGATGGGGGGCAGGACTAATAC  
TGAGTGATTGCAGAGTGCTTTATAAATATCACCTTATTTTATCGAAACCATCTGTGAACTTTCACTGAGGAAA  
AGGCCTTGACGCGGTAGAAGAGGTTGAGTCAAGGCCGGGCGCGGTGGCTCACGCCTGTAATCCAGCACTTTGGG  
AGGCCGAGGCGGTGGATCACGAGATCAGGAGATCGAGACCACCTGGCTAACACGGTGAAACCCCGTCTCTACT  
AAAAAATACAAAAGTTAGCCGGGCGTGGTGGTGGGTGCTGTAGTCCCAGCTACTCGGGAGGCTGAGGCAGGA  
GAATGGTGCGAACCCGGGAGGCGGAGCTTGAGTGAGCCAGATGGCGCCACTGCACTCCAGCCTGAGTGACAGA  
GCGAGACTCTGTCTCCA

181/615

**FIGURE 180**

MAQAVWSRLGRILWLACLLPWAPAGVAAGLYELNLTTDSPATTGAVVTISASLVAKDNGSLALPADAHLYRFHWI  
HTPLVLTGKMEKGLSSTIRVVGHVPGEFPVSVWVTAADCWMCQPVARGFVVLPITEFLVGD LVVTQNTSLPWPSS  
YLTKTVLKVSFLLHDPSNFKTALFLYSWDFGDGTQMVTEDSVVYYNYSIIIGTFTVKLKVVAEWEEVEPDATRAV  
KQKTGDFSASLKLQETLRGIQVLGPTLIQTTFQKMTVTLNFLGSPPLTVCWRLKPECLPLEEGECHPVSVASTAYN  
LTHTFRDPGDYCF SIRAENIISKTHQYHKIQVWPSRIQPAVFAPPCATLITVMLAFIMYMTLRNATQQKDMVENP  
EPPSGVRCCQMCCGPFLLLETPSEYLEIVRENHGLLPPLYKSVKTYTV

**Important features of the protein:****Signal peptide:**

amino acids 1-24

**Transmembrane domain:**

amino acids 339-362

**N-glycosylation sites.**

amino acids 34-37, 58-61, 142-145, 197-200, 300-303 and 364-367



182/615

**FIGURE 181**

CGGACGCGTGGGCGGCGGCTGCGGAACTCCCGTGGAGGGGCGGCTGGGCCCTCGGGCCTGACAG**ATGG**CAGTGGC  
CACTGCGGCGGCGAGTACTGGCCGCTCTGGGCGGGGCGCTGTGGCTGGCGGCCCGCCGGTTCTGTGGGGCCAGGGT  
CCAGCGGCTGCGCAGAGGCGGGGACCCCGGCCTCATGCACGGGAAGACTGTGCTGATCACCGGGGCGAACAGCGG  
CCTGGGCGCGCCACGGCCGCGGAGCTACTGCGCCTGGGAGCGCGGGTGATCATGGGCTGCCGGGACCGCGCGG  
CGCCGAGGAGGCGGCGGGTCAGCTCCGCGCGAGCTCCGCCAGGCCGCGGAGTGCGGGCCAGAGCCTGGCGTCAG  
CGGGGTGGGCGAGCTCATAGTCCGGGAGCTGGACCTCGCCTCGCTGCGCTCGGTGCGCGCCTTCTGCCAGGAAAT  
GCTCCAGGAAGAGCCTAGGCTGGATGTCTTGATCAATAACGCAGGGATCTTCCAGTGCCCTTACATGAAGACTGA  
AGATGGGTTTGAGATGCAGTTCGGAGTGAACCATCTGGGGCACTTTCTACTCACCAATCTTCTCCTTGGACTCCT  
CAAAAGTTCAGCTCCAGCAGGATTGTGGTAGTTTCTTCCAACTTTATAAATACGGAGACATCAATTTTGATGA  
CTTGAACAGTGAACAAAGCTATAATAAAAGCTTTTGTTATAGCCGGAGCAAAGTGGCTAACATTCTTTTACCAG  
GGAAGTACCCCGCGCTTAGAAGGCACAAATGTCACCGTCAATGTGTTGCATCCTGGTATTGTACGGACAAATCT  
GGGAGGCACATACACATTCCTACTGTTGGTCAAACCACTCTTCAATTTGGTGTGATGGGCTTTTTTCAAACTCC  
AGTAGAAGGTGCCAGACTTCCATTTATTTGGCCTCTTACCTGAGGTAGAAGGAGTGTGAGGAAGATACTTTGG  
GGATTGTAAAGAGGAAGAACTGTTGCCCAAAGCTATGGATGAATCTGTTGCAAGAAAAGTCTGGGATATCAGTGA  
AGTGATGGTTGGCCTGCTAAAA**TAGGA**ACAAGGAGTAAAGAGCTGTTTATAAACTGCATATCAGTTATATCTG  
TGATCAGGAATGGTGTGGATTGAGAACTTGTTACTTGAAGAAAAAGAAATTTTGATATTGGAATAGCCTGCTAAGA  
GGTACATGTGGGTATTTTGGAGTTACTGAAAAATTATTTTGGGATAAGAGAATTTAGCAAGAGATGTTTTAAAT  
ATATATAGTAAGTATAATGAATAATAAGTACAATGAAAAATACAATTATATTGTAATAATTATACTGGGCAAGCA  
TGGATGACATATTAATATTTGTCAGAATTAAAGTACTCAAAGTGCTATCGAGAGGTTTTTCAAGTATCTTTGAGT  
TTCATGGCCAAAGTGTTAACTAGTTTTACTACAATGTTTGGTGTGTTGTGTGGAATTTATCTGCCTGGTGTGTGCA  
CACAAGTCTTACTTGAATAAAATTTACTGGTAC

183/615

**FIGURE 182**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58747  
<subunit 1 of 1, 336 aa, 1 stop  
<MW: 36865, pI: 9.15, NX(S/T): 2  
MAVATAAAVLAAALGGALWLAARRFVGPRVQRLRRGGDPGLMHGKTVLITGANSGLGRATAAELLRLGARVIMGCR  
DRARAEEAAGQLRREL RQAAECGPEPGVSGVGELIVRELDLASLRSVRAFCQEMLQEEPRLDVLINNAGIFQCPY  
MKTEDGFEMQFGVNH LGHFLLTNLLGLLKSSAPSRIVVVSSKLYKYGDINFDDLNSEQSYNKSFCYSRSLANI  
LFTRELARRLEGTNVTNVNLHPGIVRTNLGRHIHIPLLVKPLFNLVSWAFFKTPVEGAQTSIYLASSPEVEGVSG  
RYFGDCKEEELLPKAMDESVARKLWDISEVMVGLLK

**Important features:****Signal peptide:**

amino acids 1-21

**Short-chain alcohol dehydrogenase family protein**

amino acids 134-144, 44-56 and 239-248

**N-glycosylation site.**

amino acids 212-215 and 239-242

184/615

**FIGURE 183**

AACAGGATCTCCTCTTGCAGTCTGCAGCCCAGGACGCTGATTCCAGCAGCGCCTTACCGCGCAGCCCGAAGATTC  
ACT**ATGGT**GAAAATCGCCTTCAATACCCCTACCGCCGTGCAAAAGGAGGAGGCGCGGCAAGACGTGGAGGCCCTC  
CTGAGCCGCACGGTCAGAACTCAGATACTGACCGGCAAGGAGCTCCGAGTTGCCACCCAGGAAAAAGAGGGCTCC  
TCTGGGAGATGTATGCTTACTCTCTTAGGCCTTTTCATTCATCTTGGCAGGACTTATTGTTGGTGGAGCCTGCATT  
TACAAGTACTTCATGCCCAAGAGCACCATTTACCGTGGAGAGATGTGCTTTTTTGGATTCTGAGGATCCTGCAAAAT  
TCCCTTCGTGGAGGAGAGCCTAACTTCCTGCCTGTGACTGAGGAGGCTGACATTCGTGAGGATGACAACATTGCA  
ATCATTGATGTGCCTGTCCCCAGTTTCTCTGATAGTGACCCTGCAGCAATTATTCATGACTTTGAAAAGGGAATG  
ACTGCTTACCTGGACTTGTTGCTGGGGAACTGCTATCTGATGCCCCCTCAATACTTCTATTGTTATGCCTCCAAAA  
AATCTGGTAGAGCTCTTTGGCAAACCTGGCGAGTGGCAGATATCTGCCTCAAACCTTATGTGGTTCGAGAAGACCTA  
GTTGCTGTGGAGGAAATTCGTGATGTTAGTAACCTTGGCATCTTTATTTACCAACTTTGCAATAACAGAAAGTCC  
TTCCGCCTTCGTGCGCAGAGACCTCTTGCTGGGTTTCAACAAACGTGCCATTGATAAATGCTGGAAGATTAGACAC  
TTCCCCAACGAATTTATTGTTGAGACCAAGATCTGTCAAGAG**TAA**GAGGCAACAGATAGAGTGTCTTGTTAATA  
AGAAGTCAGAGATTTACAATATGACTTTAACATTAAGGTTTATGGGATACTCAAGATATTTACTCATGCATTTAC  
TCTATTGCTTATGCTTTAAAAAAGGAAAAAATACTACTAACCCTGCAAGCTCTTGTCAAATTTTAGTT  
TAATTGGCATTGCTTGTTTTTGAACCTGAAATTACATGAGTTTCATTTTTCTTTGCATTTATAGGGTTTAGAT  
TTCTGAAAGCAGCATGAATATATCACCTAACATCCTGACAATAAATCCATCCGTTGTTTTTTTGTGTTGTTGT  
TTTTCTTTTCCCTTTAAGTAAGCTCTTTATTCATCTTATGGTGGAGCAATTTTAAATTTGAAATATTTTAAATT  
GTTTTTGAACCTTTTTGTGTAAATATATCAGATCTCAACATTGTTGGTTCTTTTGTGTTTTTCATTTGTACAAC  
TTCTTGAATTTAGAAATTACATCTTGCAGTTCTGTTAGGTGCTCTGTAATTAACCTGACTTATATGTGAACAAT  
TTTCATGAGACAGTCATTTTAACTAATGCAGTGATTCTTCTCACTACTATCTGTATTGTGGAATGCACAAAAT  
TGTGTAGGTGCTGAATGCTGTAAGGAGTTTAGGTTGTATGAATTCTACAACCCTATAATAAATTTTACTCTATAC  
AAAAAAAAAAAAAAAAAAAA

185/615

**FIGURE 184**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58828  
<subunit 1 of 1, 263 aa, 1 stop  
<MW: 29741, pI: 5.74, NX(S/T): 1  
MVKIAFNTPTAVQKEEARQDVEALLSRTVTRTQILTGKELRVATQKEGSSGRCMLTLLGLSFILAGLIVGGACIY  
KYFMPKSTIYRGEMCFFDSEDPANSLRGGEFNFPLPVTEEADIREDDNIAIIDVPVPSFSDSDPAATIHDFEKGMT  
AYLDLLLGNCYLMPLNTSIVMPPKNLVELFGKLASGRYLPQTYVVREDLVAVEEIRDVSNLGFITYQLCNNRKSF  
RLRRRDLLLGFNKRAIDKCWKIRHFPNEFIVETKICQE

**Type II transmembrane domain:**  
amino acids 53-75

**N-glycosylation site.**  
amino acids 166-170

**Casein kinase II phosphorylation site.**  
amino acids 35-39, 132-136, 134-138

**N-myristoylation site.**  
amino acids 66-72, 103-109

**Prokaryotic membrane lipoprotein lipid attachment site.**  
amino acids 63-74

186/615

**FIGURE 185**

GCTCAAGACCCAGCAGTGGGACAGCCAGACAGACGGCACGATGGCACTGAGCTCCCAGATCTGGGCCGCTTGCCT  
CCTGCTCCTCCTCCTCCTCGCCAGCCTGACCAGTGGCTCTGTTTTCCACAACAGACGGGACAACTTGCAGAGCT  
GCAACCCAGGACAGAGCTGGAGCCAGGGCCAGCTGGATGCCCATGTTCCAGAGGCGAAGGAGGCGAGACACCCA  
CTTCCCCATCTGCATTTTCTGCTGCGGCTGCTGTCATCGATCAAAGTGTGGGATGTGCTGCAAGACGTAGAACCT  
ACCTGCCCTGCCCCCGTCCCCTCCCTTCCTTATTTATTCCTGCTGCCCCAGAACATAGGTCTTGGAATAAAATGG  
CTGGTTCTTTTGTTCCTTCCAA  
AA

187/615

**FIGURE 186**

MALSSQIWAACLLLLLLLASLTSGSVFPQQTGQLAELQPQDRAGARASWMPMFQRRRRRDTHFPICIFCCGCCHR  
SKCGMCCKT

188/615

**FIGURE 187**

CTGTCAGGAAGGACCATCTGAAGGCTGCAATTTGTTCTTAGGGAGGCAGGTGCTGGCCTGGCCTGGATCTTCCAC  
CATGTTCTCTGTTGCTGCCTTTTGATAGCCTGATTGTCAACCTTCTGGGCATCTCCCTGACTGTCTCTTACCCCT  
CCTTCTCGTTTTCATCATAGTGCCAGCCATTTTGGAGTCTCCTTTGGTATCCGCAAACCTCTACATGAAAAGTCT  
GTTAAAAATCTTTGCGTGGGCTACCTTGAGAATGGAGCGAGGAGCCAAGGAGAAGAACCACCAGCTTTACAAGCC  
CTACACCAACGGAATCATTGCAAGGATCCCACTTCACTAGAAGAAGAGATCAAAGAGATTCTGCGAAGTGGTAG  
TAGTAAGGCTCTGGACAACACTCCAGAGTTCGAGCTCTCTGACATTTTCTACTTTTGCCGAAAGGAATGGAGAC  
CATTATGGATGATGAGGTGACAAAGAGATTCTCAGCAGAAGAACTGGAGTCTTGAACCTGCTGAGCAGAACCAA  
TTATAACTTCCAGTACATCAGCCTTCGGCTCACGGTCTGTGGGGGTAGGAGTGTGATTGCGTACTGCTTTCT  
GCTGCCGCTCAGGATAGCACTGGCTTTCACAGGGATTAGCCTTCTGGTGGTGGGCACAACCTGTGGTGGGATACTT  
GCCAATGGGAGGTTTAAAGGAATTCATGAGTAAACATGTTCACTTAATGTGTTACCGGATCTGCGTGGCAGCGCT  
GACAGCCATCATCACCTACCATGACAGGGAAAACAGACCAAGAAATGGTGGCATCTGTGTGGCCAATCATACCTC  
ACCGATCGATGTGATCATCTTGGCCAGCGATGGCTATTATGCCATGGTGGGTCAAGTGACCGGGGGACTCATGGG  
TGTGATTTCAGAGAGCCATGGTGAAGGCCTGCCACACGCTCTGGTTTGGAGCGCTCGGAAGTGAAGGATCGCCACCT  
GGTGGCTAAGAGACTGACTGAACATGTGCAAGATAAAAGCAAGCTGCCTATCCTCATCTTCCCAGAAGGAACCTG  
CATCAATAATACATCGGTGATGATGTTCAAAAAGGGAAGTTTTGAAATTGGAGCCACAGTTTACCCTGTTGCTAT  
CAAGTATGACCCTCAATTTGGCGATGCCTTCTGGAACAGCAGCAAATACGGGATGGTGACGTACCTGCTGCGAAT  
GATGACCAGCTGGGCCATTGTCTGCAGCGTGTGGTACCTGCCTCCCATGACTAGAGAGGCAGATGAAGATGCTGT  
CCAGTTTGCGAATAGGGTGAAATCTGCCATTGCCAGGCAGGGAGGACTTGTGGACCTGCTGTGGGATGGGGGCT  
GAAGAGGGAGAAGGTGAAGGACACGTTCAGGAGGAGCAGCAGAAGCTGTACAGCAAGATGATCGTGGGGAACCA  
CAAGGACAGGAGCCGCTCTGAGCCTGCCTCCAGCTGGCTGGGGCCACCGTGCGGGGTGCCAACGGGCTCAGAGC  
TGGAGTTGCCGCCGCCGCCCTTCTGCTGTGTCTTCCAGACTCCAGGGCTCCCCGGGCTGCTCTGGATCCCAG  
GACTCCGGCTTTCCGCGAGCCGCGAGCGGATCCCTGTGCACCCGGCGCAGCCTACCTTGGTGGTCTAAACGGAT  
GCTGCTGGGTGTTGCGACCCAGGACGAGATGCCTTGTCTTTTACAATAAGTCGTTGGAGGAATGCCATTAAAG  
TGAACCTCCACCTTTGCACGCTGTGCGGGCTGAGTGGTTGGGGAGATGTGGCCATGGTCTTGTGCTAGAGATGG  
CGGTACAAGAGTCTGTTATGCAAGCCGCTGTGCCAGGGATGTGCTGGGGCGGCCACCCGCTCTCCAGGAAAGGC  
ACAGCTGAGGCACTGTGGCTGGCTTCGGCCTCAACATCGCCCCCAGCCTTGGAGCTCTGCAGACATGATAGGAAG  
GAAACTGTCTCTGCAGGGGCTTTCAGCAAAATGAAGGGTTAGATTTTATGCTGCTGTGATGGGGTTACTAAA  
GGGAGGGGAAGAGGCCAGGTGGGCCGCTGACTGGGCCATGGGGAGAACGTGTGTTGCTACTCCAGGCTAACCCCTG  
AACTCCCCATGTGATGCGCGCTTTGTTGAATGTGTCTCGGTTTCCCCATCTGTAATATGAGTCGGGGGGAATG  
GTGGTGATTCTACCTCACAGGGCTGTTGTGGGGATTAAAGTGCTGCGGGTGAGTGAAGGACACATCACGTTTCA  
TGTTTCAAGTACAGGCCACAAAACGGGGCACGGCAGGCCTGAGCTCAGAGCTGCTGCACTGGGCTTTGGATTTG  
TTCTTGTGAGTAAATAAACTGGCTGGTGAATGA

189/615

**FIGURE 188**

MFLLLPFDSLIVNLLGISLTVLFTLLLVFIIIVPAIFGVSEFGIRKLYMKSLLKIFAWATLRMERGAKEKNHQLYKP  
YTNGIIAKDPTSLEEEIKEIRRGSSKALDNTPEFELSDIFYFCRKGMETIMDDEVTKRFSAAELESWNLLSRTN  
YNFQYISLRLTVLWGLGVLIRYCFLPLRLIALAFTGISLLVVGTTVVGYLPNGRFKEFMSKHVHLMCYRICVRAL  
TAIITYHDRENRRPRNGGICVANHTSPIDVILLASDGYAMVGQVHGGLMGVIQRAMVKACPHVWFERSEVKDRHL  
VAKRLTEHVQDKSKLPILIFPEGTCINNTSVMFKKGSFEIGATVYPVAIKYDPQFGDAFWNSSKYGMVTYLLRM  
MTSWAIVCSVWYLPMTREADEDAVQFANRVKSAIARQGGLVDLLWDGGLKREKVKDTFKEEQQKLYSKMIVGNH  
KDRSRS



190/615

**FIGURE 189**

GCCCCTCGAAACCAGGACTCCAGCACCTCTGGTCCCGCCCTCACCCGGACCCCTGGCCCTCACGTCTCCTCCAGG  
GATGCGCGCTGGCGGCTTTGATGATCGCCCTCGGCAGCCTCGGCCTCCACACCTGGCAGGCCAGGCTGTTCCCAC  
CATCCTGCCCCCTGGGCTGGCTCCAGACACCTTTGACGATACCTATGTGGGTTGTGCAGAGGAGATGGAGGAGAA  
GGCAGCCCCCTGCTAAAGGAGGAAATGGCCCACCATGCCCTGCTGCGGGAATCCTGGGAGGCAGCCAGGAGAC  
CTGGGAGGACAAGCGTCCAGGGCTTACCTTGCCCCCTGGCTTCAAAGCCCAGAATGGAATAGCCATTATGGTCTA  
CACCAACTCATCGAACACCTTGTACTGGGAGTTGAATCAGGCCGTGCGGACGGGCGGAGGCTCCCGGGAGCTCTA  
CATGAGGCACTTTCCCTTCAAGGCCCTGCATTTCTACCTGATCCGGGCCCTGCAGCTGCTGCGAGGCAGTGGGGG  
CTGCAGCAGGGGACCTGGGGAGGTGGTGTTCGAGGTGTGGGCAGCCTTCGCTTTGAACCCAAGAGGCTGGGGGA  
CTCTGTCCGCTTGGGCCAGTTTGCCTCCAGCTCCCTGGATAAGGCAGTGGCCACAGATTTGGGGAGAAGAGGCG  
GGGCTGTGTCTGCGCCAGGGGTGCAGCTAGGGTCACAATCTGAGGGGGCCTCCTCTCTGCCCCCTGGAAGAC  
TCTGCTCTTGGCCCCTGGAGAGTTCCAGCTCTCAGGGGTGGGGCCCTGAAAGTCCAACATCTGCCACTTAGGAGC  
CCTGGGAACGGGTGACCTTCATATGACGAAGAGGCACCTCCAGCAGCCTTGAGAAGCAAGAACATGGTTCCGGAC  
CCAGCCCTAGCAGCCTTCTCCCCAACCAGGATGTTGGCTGGGGAGGCCACAGCAGGGCTGAGGGAACTCTGCTA  
TGTGATGGGGACTTCCTGGGACAAGCAAGGAAAGTACTGAGGCAGCCACTTGATTGAACGGTGTGCAATGTGGA  
GACATGGAGTTTATTGAGGTAGCTACGTGATTAAATGGTATTGCAGTGTGA

191/615

**FIGURE 190**

MALAALMIALGSLGLHTWQAQAVPTILPLGLAPDTFDDTYVGCAEEMEEKAAPLLKEEMAHHALLRESWEAAQET  
WEDKRRGLTLPPGFKAQNGIAIMVYTNSSNTLYWELNQAVRTGGGSRELYMRHFPPKALHFYLIRALQLLRGSGG  
CSRGPGEVVFRGVGSLRFEPKRLGDSVRLGQFASSSLDKAVAHRFGEKRRGCVSAPGVQLGSQSEGASSLPWKT  
LLLAPGEFQLSGVGP

192/615

**FIGURE 191**

GTGGCTTCATTTCACTGGCTGACTTCCAGAGAGCAATATGGCTGGTTCCCCAACATGCCTCACCCCTCATCTATAT  
CCTTTGGCAGCTCACAGGGTCAGCAGCCTCTGGACCCGTGAAAGAGCTGGTCGGTTCCGTTGGTGGGGCCCGTGAC  
TTTCCCCCTGAAGTCCAAAGTAAAGCAAGTTGACTCTATTGTCTGGACCTTCAACACAACCCCTCTTGTCACCAT  
ACAGCCAGAAGGGGGCACTATCATAGTGACCCAAAATCGTAATAGGGAGAGAGTAGACTTCCCAGATGGAGGCTA  
CTCCCTGAAGCTCAGCAAACCTGAAGAAGAATGACTCAGGGATCTACTATGTGGGGATATACAGCTCATCACTCCA  
GCAGCCCTCCACCCAGGAGTACGTGCTGCATGTCTACGAGCACCTGTCAAAGCCTAAAGTCACCATGGGTCTGCA  
GAGCAATAAGAATGGCACCTGTGTGACCAATCTGACATGCTGCATGGAACATGGGGAAGAGGATGTGATTTATAC  
CTGGAAGGCCCTGGGGCAAGCAGCCAATGAGTCCCATAATGGGTCCATCCTCCCCATCTCCTGGAGATGGGGAGA  
AAGTGATATGACCTTCATCTGCGTTGCCAGGAACCCTGTCAGCAGAACTTCTCAAGCCCCATCCTTGCCAGGAA  
GCTCTGTGAAGGTGCTGCTGATGACCCAGATTCTCCATGGTCCTCCTGTGTCTCCTGTTGGTGCCCCCTCCTGCT  
CAGTCTCTTTGTACTGGGGCTATTTCTTTGGTTTCTGAAGAGAGAGAGACAAGAAGAGTACATTGAAGAGAAGAA  
GAGAGTGGACATTTGTGGGAAACTCCTAACATATGCCCCATTCTGGAGAGAACACAGAGTACGACACAATCCC  
TCACACTAATAGAACAATCCTAAAGGAAGATCCAGCAAATACGGTTTACTCCACTGTGGAAATACCGAAAAAGAT  
GGAAAATCCCCACTCACTGCTCACGATGCCAGACACACCAAGGCTATTTGCCTATGAGAATGTTATCTAGACAGC  
AGTGCACTCCCCTAAGTCTCTGCTCA

193/615

**FIGURE 192**

MAGSPTCLTLIYILWQLTGSAAAGPVKELVGSVGGAVTFPLKSKVKQVDSIVWTFNTTPLVTIQPEGGTIIIVTON  
RNRERVDFPDGGYSLKLSKLKKNDSGIYYVGIYSSSLQQPSTQEYVLHVYEHLSKPKVTMGLQSNKNGTCVTNLT  
CCMEHGEEEDVIYTWKALGQAANESHNGSILPISWRWGESDMTFICVARNPVSRNFSSPILARKLCEGAADDPDSS  
MVLCLLLVPLLLSLFVLGLFLWFLKRERQEEYIEKKRVDICRETPNICPHSGENTEYDTIPHTNRTILKEDPA  
NTVYSTVEIPKKMENPHSLTMDTPRLFAYENVI

FIGURE 193

[illegible]

195/615

**FIGURE 194**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58852
><subunit 1 of 1, 283 aa, 1 stop
><MW: 29191, pI: 4.52, NX(S/T): 0
MVSAAAPSLILLILLLLLSVLPATDARSVPLKATFLEDVAGSGEAGSSASSPSLPPPWTLPALSPTSMGPQPTTLG
GPSPTNFDLGIVDFFRQYVMLIAVVGSLAFLLMFIVCAAVITRQKQKASAYYPSSFPPKKYVDQSDRAGGPRAF
SEVPDRAPDSRPEEALDSSRQLQADILAATQNLKSPTRAALGGGDGARMVEGRGAEEEEKGSQEGDQEVQGHGVP
VETPEAQEEPCSGVLEGAVVAGEGQGELEGSLLLAQEAQGPVGPPEPCACSSVHPSV
```

**Signal peptide:**  
amino acids 1-25

**Transmembrane domain:**  
amino acids 94-118

**N-myristoylation site.**  
amino acids 18-24, 40-46, 46-52, 145-151, 192-198, 193-199, 211-217, 238-244,  
242-248

196/615

**FIGURE 195**

GAAAGACGTGGTCCTGACAGACAGACAATCCTATTCCCTACCAAAATGAAGATGCTGCTGCTGCTGTGTTTGGGA  
CTGACCCTAGTCTGTGTCCATGCAGAAGAAGCTAGTTCTACGGGAAGGAACTTAATGTAGAAAAGATTAATGGG  
GAATGGCATACTATTATCCTGGCCTCTGACAAAAGAGAAAAGATAGAAGAACATGGCAACTTTAGACTTTTTCTG  
GAGCAAATCCATGTCTTGGAGAATTCCCTAGTTCTTAAAGTCCATACTGTAAGAGATGAAGAGTGCTCCGAATTA  
TCTATGGTTGCTGACAAAACAGAAAAGGCTGGTGAATATTCTGTGACGTATGATGGATTCAATACATTTACTATA  
CCTAAGACAGACTATGATAACTTTCTTATGGCTCACCTCATTAACGAAAAGGATGGGGAAACCTTCCAGCTGATG  
GGGCTCTATGGCCGAGAACCAGATTTGAGTTCAGACATCAAGGAAAGGTTTGCACAACTATGTGAGGAGCATGGA  
ATCCTTAGAGAAAATATCATTGACCTATCCAATGCCAATCGCTGCCTCCAGGCCCGAGAAATGAAGAATGGCCTGA  
GCCTCCAGTGTTGAGTGGACACTTCTCACCAGGACTCCACCATCATCCCTTCCTATCCATACAGCATCCCCAGTA  
TAAATTCTGTGATCTGCATTCCATCCTGTCTCACTGAGAAGTCCAATTCCAGTCTATCAACATGTTACCTAGGAT  
ACCTCATCAAGAATCAAAGACTTCTTTAAATTTCTCTTTGATACACCCTTGACAATTTTTTCATGAAATTATTCCT  
CTTCTGTTCATAAATGATTACCCTTGCACTTAA

197/615

**FIGURE 196**

MKMLLLLCGLTLVCHAEFASSTGRNFNVEKINGEWHTIILASDKREKIEEHG NFR LFLEQIHVLENSLV LKVH  
TVRDEECSELSMVADKTEKAGEYSVTYDGFNTFTIPKTDYDNFLMAHLINEKDGETFQLMGLYGREPDLSSDIKE  
RFAQLCEEHGILRENIIDLSNANRCLQARE



198/615

**FIGURE 197**

GGCTCGAGCGTTTCTGAGCCAGGGGTGACCATGACCTGCTGCGAAGGATGGACATCCTGCAATGGATTGAGCCTG  
CTGGTTCTACTGCTGTTAGGAGTAGTTCTCAATGCGATACCTCTAATTGTCAGCTTAGTTGAGGAAGACCAATTT  
TCTCAAACCCCATCTCTTGCTTTGAGTGGTGGTCCCAGGAATTATAGGAGCAGGTCTGATGGCCATTCCAGCA  
ACAACAATGTCCTTGACAGCAAGAAAAAGAGCGTGCTGCAACAACAGAACTGGAATGTTTCTTTCATCATTTTTC  
AGTGTGATCACAGTCATTGGTGCTCTGTATTGCATGCTGATATCCATCCAGGCTCTCTTAAAGGTCCTCTCATG  
TGTAATTCTCCAAGCAACAGTAATGCCAATTGTGAATTTTCATTGAAAAACATCAGTGACATTATCCAGAATCC  
TTCAACTTGCAGTGGTTTTTCAATGACTCTTGTGCACCTCCTACTGGTTTCAATAAACCCACCAGTAACGACACC  
ATGGCGAGTGGCTGGAGAGCATCTAGTTTCCACTTCGATTCTGAAGAAAACAAACATAGGCTTATCCACTTCTCA  
GTATTTTTAGGTCTATTGCTTGTTGGAATTCTGGAGGTCCTGTTGGGCTCAGTCAGATAGTCATCGGTTTCCTT  
GGCTGTCTGTGTGGAGTCTCTAAGCGAAGAAGTCAAATTGTGTAGTTTAATGGGAATAAAATGTAAGTATCAGTA  
GTTTGAAAAAAAAAAAA

199/615

**FIGURE 198**

MTCCEGWTS CNGFSLLVLLLLGVVLNAIPLIVSLVEEDQFSQNPISCFEWWFPGIIGAGLMAIPATTMSLTARKR  
ACCNRTGMFLSSFFSVITVIGALYCM LISIQALLKGPLMCNSPSNSNANCEFSLKNISDIHPESFNLOWFFNDS  
CAPPTGFNKPTSNDT MASGWRASSFHFDSEENKHRLIHFSVFLGLLL VGILEVLFGLSQIVIGFLGCLCGVSKRR  
SQIV

200/615

**FIGURE 199**

ATCCGTTCTCTGCGCTGCCAGCTCAGGTGAGCCCTCGCCAAGGTGACCTCGCAGGACACTGGTGAAGGAGCAGTG  
AGGAACCTGCAGAGTCACACAGTTGCTGACCAATTGAGCTGTGAGCCTGGAGCAGATCCGTGGGCTGCAGACCCC  
CGCCCCAGTGCCTCTCCCCCTGCAGCCCTGCCCCCTCGAACTGTGACATGGAGAGAGTGACCCCTGGCCCTTCTCCT  
ACTGGCAGGCCTGACTGCCTTGGAAGCCAATGACCCATTTGCCAATAAAGACGATCCCTTCTACTATGACTGGAA  
AAACCTGCAGCTGAGCGGACTGATCTGCGGAGGGCTCCTGGCCATTGCTGGGATCGCGGCAGTTCTGAGTGGCAA  
ATGCAAATACAAGAGCAGCCAGAAGCAGCACAGTCCTGTACCTGAGAAGGCCATCCCACTCATCACTCCAGGCTC  
TGCCACTACTTGCTTGAGCACAGGACTGGCCTCCAGGGATGGCCTGAAGCCTAACACTGGCCCCCAGCACCTCCTC  
CCCTGGGAGGCCTTATCCTCAAGGAAGGACTTCTCTCCAAGGGCAGGCTGTTAGGCCCTTTCTGATCAGGAGGC  
TTCTTTATGAATTAAACTCGCCCCACCACCCCTCA

201/615

**FIGURE 200**

MERVTLALLLLAGLTALEANDPFANKDDPFYYDWKNLQLSGLICGGLLAIAAGIAAVLSGKCKYKSSQKQHSPVPE  
KAIP LITPGSATTC

FIGURE 201

[illegible]

**FIGURE 202**

amino acids 301-304

204/615

**FIGURE 203**

GGAGAAGAGGTTGTGTGGGACAAGCTGCTCCCGACAGAAGGATGTCGCTGCTGAGCCTGCCCTGGCTGGGCCTCA  
GACCGGTGGCAATGTCCCATGGCTACTCCTGCTGCTGGTTGTGGGCTCCTGGCTACTCGCCCGCATCCTGGCTT  
GGACCTATGCCTTCTATAACAAGTGGCGCCGGCTCCAGTGTTCACAGCCCCAAAACGGAAGTGGTTTTGGG  
GTCACCTGGGCCTGATCACTCCTACAGAGGAGGGCTTGAAGGACTCGACCCAGATGTCGGCCACCTATTCCCAGG  
GCTTTACGGTATGGCTGGGTCCCATCATCCCCTTCATCGTTTTATGCCACCCTGACACCATCCGGTCTATCACCA  
ATGCCTCAGCTGCCATTGCACCCAAGGATAATCTCTTCATCAGGTTCCCTGAAGCCCTGGCTGGGAGAAGGGATAC  
TGCTGAGTGGCGGTGACAAGTGGAGCCGCCACCGTCGGATGCTGACGCCCGCCTTCCATTTCAACATCCTGAAGT  
CCTATATAACGATCTTCAACAAGAGTGCAAACATCATGCTTGACAAGTGGCAGCACCTGGCCTCAGAGGGCAGCA  
GTCTCTGGACATGTTTGAGCACATCAGCCTCATGACCTTGGACAGTCTACAGAAATGCATCTTCAGCTTTGACA  
GCCATTGTCAGGAGAGGCCCCAGTGAATATATTGCCACCATCTTGGAGCTCAGTGCCCTTGTAGAGAAAAGAAGCC  
AGCATATCCTCCAGCACATGGACTTTCTGTATTACCTCTCCCATGACGGGCGGCGCTTCCACAGGGCCTGCCGCC  
TGGTGCATGACTTCACAGACGCTGTATCCGGGAGCGGCGTCGCACCCTCCCCACTCAGGGTATTGATGATTTTT  
TCAAAGACAAAGCCAAAGTCCAAGACTTTGGATTTCATTGATGTGCTTCTGCTGAGCAAGGATGAAGATGGGAAGG  
CATTGTGATGAGGATATAAGAGCAGAGGCTGACACCTTCATGTTTGGAGGCCATGACACCACGGCCAGTGGCC  
TCTCCTGGGTCCTGTACAACCTTGCGAGGCACCCAGAATACCAGGAGCGCTGCCGACAGGAGGTGCAAGAGCTTC  
TGAAGGACCGCGATCCTAAAGAGATTGAATGGGACGACCTGGCCCAGCTGCCCTTCTGACCATGTGCGTGAAGG  
AGAGCCTGAGGTTACATCCCCAGCTCCCTTCATCTCCCGATGCTGCACCCAGGACATTGTTCTCCAGATGGCC  
GAGTCATCCCCAAAGGCATTACCTGCCTCATCGATATTATAGGGGTCCATCACAACCCAACTGTGTGGCCGGATC  
CTGAGGTCTACGACCCCTTCGCTTTGACCCAGAGAACAGCAAGGGGAGGTACCTCTGGCTTTTATTCCTTTCT  
CCGAGGGCCCAGGAACTGCATCGGGCAGGCGTTCGCCATGGCGGAGATGAAAGTGGTCTGGCGTTGATGCTGC  
TGCACTTCGGGTTCTGCCAGACCACTGAGCCCCGAGGAAGCTGGAATTGATCATGCGCGCCGAGGGCGGGC  
TTTGGCTGCGGTTGGAGCCCCCTGAATGTAGGCTTGCAGTGACTTTTCTGACCCATCCACCTGTTTTTTTGCAGATT  
GTCATGAATAAAACGGTGCTGTCAA

205/615

**FIGURE 204**

MSLLSLPWLGLRPVAMSPWLLLLLVVGSWLLARILAWTYAFYNNCRRLQCFPPQPPKRNWFWGHLGLITPTEEGLK  
DSTQMSATYSQGFTVWLGPPIPFIVLCHPDTIRSITNASAAIAPKDNLFIRFLKPWLGEGILLSGGDKWSRHRM  
LTPAFHFNILKSYITIFNKSANIMLDKWQHHLASEGSSRLDMFEHISLMTLDSLQKCIFSFDSHCQERPSEYIATI  
LELSALVEKRSQHILQHMDFLYYLSHDGRRFHRACRLVHDFTDVIRERRRTLPTQGIDDFKDKAKSKTLDFID  
VLLLSKDEDGKALSDEDIRAEADTFMFGGHDTTASGLSWVLYNLARHPEYQERCQEVQELLKDRDPKEIEWDDL  
AQLPFLTMCVKESLRLHPPAPFISRCCTQDIVLPDGRVIPKGITCLIDIIGVHHNPTVWPDPEVYDPFRFDPENS  
KGRSPLAFIPFSAGPRNCIGQAFAMAEMKVVLALMLLHFRFLPDHTEPRRKLELIMRAEGGLWLRVEPLNVGLQ



206/615

**FIGURE 205**

TCCCTTGACAGGTCTGGTGGCTGGTTCGGGGTCTACTGAAGGCTGTCTTGATCAGGAACTGAAGACTCTCTGCT  
TTTGCCACAGCAGTTCCTGCAGCTTCCTTGAGGTGTGAACCCACATCCCTGCCCCAGGGCCACCTGCAGGACGC  
CGACACCTACCCCTCAGCAGACGCCGGAGAGAAATGAGTAGCAACAAAGAGCAGCGGTGACAGTGTTCGTGATC  
CTCTTTGCCCTCATCACCATCCTCATCCTCTACAGCTCCAACAGTGCCAATGAGGTCTTCCATTACGGCTCCCTG  
CGGGCCCGTAGCCGCCGACCTGTCAACCTCAAGAAGTGGAGCATCACTGACGGCTATGTCCCCATTCTCGGCAAC  
AAGACACTGCCCTCTCGGTGCCACCAGTGTGTGATTGTGACGAGCTCCAGCCACCTGCTGGGCACCAAGCTGGGC  
CCTGAGATCGAGCGGGCTGAGTGTACAATCCGCATGAATGATGACCCACCACTGGCTACTCAGCTGATGTGGGC  
AACAAGACCACCTACCGCGTCTGGGCCATTCCAGTGTGTTCCGCGTCTGAGGAGGCCCCAGGAGTTTGTCAAC  
CGGACCCCTGAAACCGTGTTCATCTTCTGGGGGCCCCGAGCAAGATGCAGAAGCCCCAGGGCAGCCTCGTGCGT  
GTGATCCAGCGAGCGGGCCTGGTGTTCCTCAACATGGAAGCATATGCCGTCTCTCCCGGCCGCATGCGGCAATTT  
GACGACCTCTTCCGGGTGAGACGGGCAAGGACAGGGAGAAGTCTCATTTCGTGGTTGAGCACAGGCTGGTTTACC  
ATGGTGATCGCGGTGGAGTTGTGTGACCAGTGCATGTCTATGGCATGGTCCCCCCTAACTACTGCAGCCAGCGG  
CCCCGCTCCAGCGCATGCCCTACCACTACTACGAGCCCAAGGGGCCGACGAATGTGTACCTACATCCAGAAT  
GAGCACAGTCGCAAGGGCAACCACCACCGCTTCATCACCAGAGAAAAGGCTCTTCTCATCGTGGGCCAGCTGTAT  
GGCATCACCTTCTCCACCCCTCCTGGACCTAGGCCACCCAGCCTGTGGGACCTCAGGAGGGTCAGAGGAGAAGC  
AGCCTCCGCCAGCCGCTAGGCCAGGGACCATCTTCTGGCCAATCAAGGCTTGCTGGAGTGTCTCCAGCCAATC  
AGGGCCTTGAGGAGGATGTATCCTCCAGCCAATCAGGGCCTGGGGAATCTGTTGGCGAATCAGGGATTTGGGAGT  
CTATGTGGTTAATCAGGGGTGTCTTCTTGTGTCAGTCAGGGTCTGCGCACAGTCAATCAGGGTAGAGGGGTATT  
TCTGAGTCAATCTGAGGCTAAGGACATGTCTTCTCCATGAGGCCTTGGTTCAGAGCCCCAGGAATGGACCCCC  
AATCACTCCCCACTCTGCTGGGATAATGGGGTCTGTCCCAAGGAGCTGGGAACCTGGTGTGCCCCCTCAATTT  
CCAGCACCAGAAAGAGAGATTGTGTGGGGGTAGAAGCTGTCTGGAGGCCCGGCCAGAGAATTTGTGGGGTTGTGG  
AGGTGTGGGGGGCGGTGGGGAGGTCCCAGAGGTGGGAGGCTGGCATCCAGGTCTTGGCTCTGCCCTGAGACCTTG  
GACAAACCTTCCCCCTCTCTGGGCACCTTCTGCCACACCAAGTTCCAGTGCGGAGTCTGAGACCTTTCCAC  
CTCCCCATAAGTGCCCTCGGGTCTGTCTCCCCGTCTGGACCTCCCAGCCACTATCCCTGTGGAAGGCTCA  
GCTCTTTGGGGGTCTGGGGTGACCTCCCCACCTCCTGAAAACCTTAGGGTATTTTTCGCGAACTCCTTCAGG  
GTTGGGGGACTCTGAAGGAAACGGGACAAAACCTTAAGCTGTTTTCTTAGCCCCTCAGCCAGCTGCCATTAGCTT  
GGCTCTTAAAGGGCCAGGCCTCCTTTTCTGCCCTCTAGCAGGGAGGTTTTCCAAGTGTGGAGGCGCCTTTGGGG  
CTGCCCCCTTGTCTGGAGTCACTGGGGGCTTCCGAGGGTCTCCCTCGACCTCTGTCTGCTGCGGATGGCTGTCTG  
GGAGCTGTATCACCTGGGTCTGTCCCTGGCTCTGTATCAGGCATTTATTAAAGCTGGGCTCAGTGGGGTGT  
GTTTGTCTCCTGCTCTTCTGGAGCCTGGAAGGAAAGGGCTTCAGGAGGAGGCTGTGAGGCTGGAGGGACAGATG  
GAGGAGGCCAGCAGCTAGCCATTGCACACTGGGTGATGGGTGGGGGCGGTGACTGCCCCAGACTTGGTTTTGTA  
ATGATTTGTACAGGAATAAACACACCTACGCTCCGGAAGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

207/615

**FIGURE 206**

MSSNKEQRS AVEFVILFALITILILYSSNSANEVFHYGSLRGRSRPVLKKWSITDGYVPILGNKTLPSRCHQCV  
IVSSSSHLGLGPEIERAECTIRMNDAPTTGYSADVGNKTTYRVVAHSSVFRVLRRPQEFVNRTPETVFIFWG  
PPSKMQKPQGS LVRVIQRAGLVFPNMEAYAVSPGRMRQFDDLFRGETGKDREKSHSWLSTGWFTMVI AVELCDHV  
HVG MVPPNYCSQRRLQRMPIHYEYEPKGPDECVTYIQNEHSRKGNNHHRFITEKRVFSSWAQLYGITFSHPST

**Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 9-31 (type II)

**N-glycosylation site.**

amino acids 64-68, 115-119

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 50-54

**Casein kinase II phosphorylation site.**

amino acids 3-7, 29-33, 53-57, 197-201

**Tyrosine kinase phosphorylation site.**

amino acids 253-262

**N-myristoylation site.**

amino acids 37-43, 114-120, 290-294

208/615

**FIGURE 207**

GTAGCGCGTCTTGGGTCTCCCGGCTGCCGCTGCTGCCGCCGCCGCTCGGGTCTGGAGCCAGGAGCGACGTCAC  
CGCC**ATGG**CAGGCATCAAAGCTTGATTAGTTTGTCTTTGGAGGAGCAATCGGACTGATGTTTTTATGCTGG  
ATGTGCCCTTCCAATATACAACAAATACTGGCCCTCTTTGTTCTATTTTTTACATCCTTTCACCTATTTCCATA  
CTGCATAGCAAGAAGATTAGTGGATGATACAGATGCTATGAGTAACGCTTGTAAGGAACCTGCCATCTTTCTTAC  
AACGGGCATTGTCGTGTAGCTTTTGGACTCCCTATTGTATTTGCCAGAGCACATCTGATTGAGTGGGGAGCTTG  
TGCATTGTTCTCACAGGAAACACAGTCATCTTGTCACTATACTAGGCTTTTTCTTGGTCTTTGGAAGCAATGA  
CGACTTCAGCTGGCAGCAGTGG**TG**AAAAAGAAATTACTGAACTATTGTCAAATGGACTTCCTGTCATTTGTTGGCC  
ATTCACGCACACAGGAGATGGGGCAGTTAATGCTGAATGGTATAGCAAGCCTCTTGGGGGTATTTTAGGTGCTCC  
CTTCTCACTTTTATTGTAAGCATACTATTTTACAGAGACTTGCTGAAGGATTAAGGATTTTCTCTTTTGGA  
AAGCTTGACTGATTTACACTTATCTATAGTATGCTTTTTTGTGGTGTCTGCTGAATTTAAATATTTATGTGTTT  
TTCCTGTTAGGTTGATTTTTTTTGGAAATCAATATGCAATGTAAACACTTTTTTAATGTAATCATTTGCATTGGT  
TAGGAATTCAGAATCCGCCGGCTCTATTACTGGTCAAGTACATCTTTCTCTTAAATTAATTTAGCTCCATTA  
TTACAAAAAATTATAAAAAATAAGTTTTCAGTCAGTCAGGATGACATCACTCCCAATGTTATGCAGACATACAGAC  
GGTTGGCATACTGTATAGACTGTATACTCAGTGCAAATATAGCTGCATTATACCTCAGAGGGGCCAAGTGTTAA  
TGCCCATGCCCTCCGTTAAGGGTTGTTGGTTTTACTGGTAGACAGATGTTTTGTGGATTGAAAATTATTTATGG  
AATTGCTACAGAGGAGTGCTTTTCTTCTCAATTGTTAGAAGAATTTATGTTAAACTTTAAGGTAAGGGTGTA  
ACATTTTTGAGATAAGGTTTTTATTTATGTTTATTATTGTTAGAGTGAGTTGCAATGTGGGAAGAAATGACATTG  
AAATTCAGTTTTTGAATCCTGTTTCTATTTATAAGTGAATTTGTGATCTCTATCAACCTTTTCATGTTTTACC  
CTGTTAAATGGACATACATGGAACCACTACTGATGAGGGACAGTTGTATGTTGCATCATATATGCCAGAAAC  
CTTCTCTGCTTCCTCCTTTTGACTTATTTGGTATGTTGTATATATTACATAAAAAATAACTTTTCAAATATAGTTT  
AATAACACTTAGAAGTGTTTACTTACCTGGAATAAATTGCTATGCCGTACATTACAGAGTGCCCCCTCCCCTGCA  
AGGCTTGCCATGATTAACAAGTAACCTTGTTAGTCTTACAGATAATTCATGCATTAACAGTTTAAAGATTTAGACC  
ATGGTAATAGTAGTTCTTATTCTCTAAGGTTATATCATATGTAATTTAAAGTATTTTTAAGACAAGTTTCTGT  
ATACCTCTGAAGTGTGTTTTGATTTGAGTTCATCATGATAGATCTGCTGTTTCTTATAAAAGGCATTTGTGTGT  
GAGTTAATGCAAAGTAGCCAAGTCCAGCTATATAGCAGCTTCAGAAACATACCTGACCAAAAAATCCCAGTAAC  
CAGGCATGATCAATTTATAGTGGTCGTTTACATCTAATAATTATCAGGACTTTTTTTCAGGAGTGGGTTATAAAAA  
CATTCAAGTTGGTCTGACAGTATTTTGTAAAGGATATTGTTTGTATGTTTATTACGTATACTTACATAAAAAAT  
ATTTGCCCATCAGCCAAAACCTCAGTAATCATGACAGCTGTCTGTTGTTTTATGAAGTTTATTTCTCAAGAAAATG  
GGAATAAATTTGGGATTTGTTTCAGCTTTTTTACTAAAGATGCCATAAGCCACAGGTTTTATTGCCCTAACTTAAGC  
CATGACTTTTAGATATGAGATGACGGGAAGCAGGACGAAATATCGGCCGTGTGGCTGGAGCCTTCCCACTGGAGGC  
TGAAAGTGGCTTGTGGTATTATAATGTTTCAAGAGGAAGGTGCAGGTACACATGAGTTAGAGAGCTGGT  
GAGACAGTTGGGAACCTTTGTGCTTGTGATCTACTGGACTTTTTTTTTTGCAGGAAGTGCATTCCTCTGGTCTTC  
CCTATTTTCTGTTCTGGATGTCAGTGCAGTGCAGTCTACTGTTTTATCCACTTGGCCACAGACTTTTTCTAACA  
GCTGCGTATTATTTCTATATACTAATTGCATTGGCAGCATTTGTGTCTTTGACCTTGATACACTAGCTTGACATAGT  
GCTGTCTCTGATTTCTAGGCTAGTTACTTGAGATATGAATTTTCCATAGAATATGCACTGATACACATTACCAT  
TCTTCTATGGAAGAAAACCTTTTGATGATGAACAATAAAGATTTTAAATATCTATTTTAAAAA

209/615

**FIGURE 208**

MAGIKALISLSFGGAIGLMFLMLGICALPIYNKYWPLFVLFFYILSPIPYCIARRLVDDTDAMSNACKELAIFLT  
T  
GIVVSAFGLPIVFARAHLEWGACALVLTGNTVIFATILGFFLVFGSNDDFSWQQW

210/615

**FIGURE 209**

CTTGACAGAGAAAGAGTCTTTTGTGCAGCACCCCTTTAAAGGGTGACTCGTCCCACCTTGTGTTCTCTCTCCTGGTGC  
AGAGTTGCAAGCAAGTTTATCAGAGTATCGCCATGAAGTTTCGTCCCCTGCCTCCTGCTGGTGACCTTGTCTGCTGCC  
TGGGGACTTTGGGTCAGGCCCCGAGGCAAAAGCAAGGAAGCACTGGGGAGGAATTCATTTCCAGACTGGAGGGA  
GAGATTCTGCACTATGCGTCCCAGCAGCTTGGGGCAAGGTGCTGGAGAAGTCTGGCTTCGCGTCGACTGCCGCA  
ACACAGACCAGACCTACTGGTGTGAGTACAGGGGGCAGCCCAGCATGTGCCAGGCTTTTGCTGCTGACCCCAAAC  
CTTACTGGAATCAAGCCCTGCAGGAGCTGAGGCGCCTTCACCATGCGTGCCAGGGGGCCCCGGTGCTTAGGCCAT  
CCGTGTGCAGGGAGGCTGGACCCAGGCCCATATGCAGCAGGTGACTTCCAGCCTCAAGGGCAGCCCAGAGCCCA  
ACCAGCAGCCTGAGGCTGGGACGCCATCTCTGAGGCCCAAGGCCACAGTGAACTCACAGAAGCAACACAGCTGG  
GAAAGGACTCGATGGAAGAGCTGGGAAAAGCCAAACCCACCACCCGACCCACAGCCAAACCTACCCAGCCTGGAC  
CCAGGCCCGGAGGGAATGAGGAAGCAAAGAAGAAGGCCTGGGAACATTGTTGGAACCCCTTCCAGGCCCTGTGCG  
CCTTTCTCATCAGCTTCTTCCGAGGGTGACAGGTGAAAGACCCCTACAGATCTGACCTCTCCCTGACAGACAACC  
ATCTCTTTTTATATTATGCCGCTTTCAATCCAACGTTCTCACACTGGAAGAAGAGAGTTTCTAATCAGATGCAAC  
GGCCCAAATTCTTGATCTGCAGCTTCTCTGAAGTTTGGAAAAGAAACCTTCCTTTCTGGAGTTTGCAGAGTTCAG  
CAATATGATAGGGAACAGGTGCTGATGGGCCCAAGAGTGACAAGCATACACAATACTATTATTCTGTAGAAGTT  
TTGCTTTGTTGATCTGAGCCTTCTATGAAAGTTTAAATATGTAACGCATTCATGAATTTCCAGTGTTTCAGTAAAT  
AGCAGCTAIGTGTGTGCAAAATAAAAGAATGATTTTCAGAAAAA

211/615

**FIGURE 210**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59602
<subunit 1 of 1, 223 aa, 1 stop
<MW: 24581, pI: 9.28, NX(S/T): 0
MKFVPCLLLVTLSCCLGTLGQAPRQKQGSGTEEFHFQTGGRDSCTMRPSSLGQGAGEVWLR
VDCRNTDQTYWCEYRGQPSMCQAFADPKPYWNQALQELRRLHHACQGAPVLRPSVCREA
GPQAHMQQVTSSLKGSPEPNQQPEAGTPSLRPKATVKLTEATQLGKDSMEELGKAKPTTR
PTAKPTQPGPRPGGNEEAKKKAWEHCWKPFQALCAFLISFFRG
```

**Important features:****Signal peptide:**

Amino acids: 1-19

**N-myristoylation sites:**

Amino acids: 38-44;51-57;194-200

**DNA photolyases class 1 proteins:**

Amino acids: 58-69

**Tyrosine kinase phosphorylation site:**

Amino acids: 64-71

**N-myristoylation sites:**

Amino acids: 38-44;51-57;194-200

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids: 4-15

212/615

**FIGURE 211**

GTGCAAGGAGCCGAGGCGAG**ATG**GGGCGTCCTGGGCGGGTCCTGCTGTGGCTGCAGCTCTGCGCACTGACCCAGGCG  
GTCTCCAAACTCTGGGTCCCCAACACGGACTTCGACGTCGCAGCCAACTGGAGCCAGAACCGGACCCCGTGCGCC  
GGCGGCGCCGTTGAGTTCCTCGCGGACAAGATGGTGTCACTCCTGGTGCAAGAAGGTCACGCCGTCTCAGACATG  
CTCCTGCCGCTGGATGGGGAACCTCGTCCTGGCTTCAGGAGCCGGATTGCGCGTCTCAGACGTGGGCTCGCACCTG  
GACTGTGGCGCGGGCGAACCTGCCGTCTTCCGCGACTCTGACCGCTTCTCCTGGCATGACCGCACCTGTGGCGCT  
CTGGGGACGAGGCACCTGGCCTCTTCTTCGTGGACGCCGAGCGCGTGCCCTGCCGCCACGACGACGTCTTCTTTC  
CGCCTAGTGCCCTCCTTCCGCGTGCGGGCTCGGCCCTGGCGCTAGCCCCGTGCGTGTCCGCAGCATCTCGGCTCTGG  
GCCGGACGTTACGCGCGACGAGGACCTGGCTGTTTTCTGCGGTCCCGCGCGGGCCGCCTACGCTTCCACGGGC  
CGGGCGCGCT**TGAG**CGTGCGCCCCGAGGACTGCGCGGACCCGTGCGGCTGCGTCTGCGGCAACGCGGAGGCGCAGC  
CGTGGATCTGCGCGGCCCTGCTCCAGCCCCT

213/615

**FIGURE 212**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59603
<subunit 1 of 1, 197 aa, 1 stop
<MW: 20832, pI: 8.74, NX(S/T): 2
MGVLGRVLLWLQLCALTQAVSKLWVPNTDFDVAANWSQNRTPCAGGAVEFPADKMVSVLV
QEGHAVSDMLLPLDGELVLASGAGFGVSDVGSHLDCGAGEPAVFRSDRFSWHDRTCCGAL
GTRHLASSSWTPSACPAATTSSFRLVPPSAWGSALALAPCVSAASRLWAGRSRATRTWL
FSWRPARAAYASTGRAR
```

**Important features:****Signal peptide:**

Amino acids 1-19

**N-glycosylation site:**

Amino acids 35-39

**Glycosaminoglycan attachment site:**

Amino acids 81-85

**N-myristoylation sites:**

Amino acids 82-88;118-124;153-159

**C-type lectin domain proteins:**

Amino acids 108-118



FIGURE 213

[illegible]

215/615

**FIGURE 214**

MGPVKQLKRMFEPTRLIATIMVLLCFALTLCSAFWWHNKGLALIFCILQSLALTWYSLSFIPFARDAVKKCFVCLA

216/615

**FIGURE 215**

GGATTTTTGTGATCCGCGATTGCTCCCACGGGCGGGACCTTTGTAAGTGGGGAGGCCCAGGACAGGCCCACCC  
TGCGGGGCGGGAGGCAGCCGGGGTGAGGGAGGTGAAGAAACCAAGACGCAGAGAGGCCAAGCCCCTTGCCTTGGG  
TCACACAGCCAAAGGAGGCAGAGCCAGAATCACAACCAGATCCAGAGGCAACAGGGACATGGCCACCTGGGACG  
AAAAGGCAGTCACCCGCAGGGCCAAGGTGGCTCCCGCTGAGAGGATGAGCAAGTTCTTAAGGCACTTCACGGTCG  
TGGGAGACGACTACCATGCCTGGAACATCAACTACAAGAAATGGGAGAATGAAGAGGAGGAGGAGGAGGAGC  
AGCCACCACCCACACCACTCTCAGGCGAGGAAGGCAGAGCTGCAGCCCCCTGACGTTGCCCTGCCCTGGCCCCG  
CAGCCAGGGCCCCCTTGACTTCAGGGGCATGTTGAGGAACTGTTGAGCTCCACAGGTTTCAGGTCATCATCA  
TCTGCTTGGTGGTTCTGGATGCCCTCCTGGTGCTTGCTGAGCTCATCCTGGACCTGAAGATCATCCAGCCCGACA  
AGAATAACTATGCTGCCATGGTATTCCACTACATGAGCATCACCATCTTGGTCTTTTTTATGATGGAGATCATCT  
TTAAATTATTTGTCTTCCGCCTGAGTTCTTTTACCACAAGTTTGAGATCCTGGATGCCCGTCGTGGTGGTGGTCT  
CATTCATCCTGGACATTGTCCTCCTGTTCCAGGAGCACCAGTTTGAGGCTCTGGGCCTGCTGATTCTGCTCCGGC  
TGTGGCGGGTGGCCCGGATCATCAATGGGATTATCATCTCAGTTAAGACACGTTTCAAGACGGCAACTCTTAAGGT  
TAAAACAGATGAATGTACAATTGGCCGCCAAGATTCAACACCTTGAGTTGAGCTGCTCTGAGAAGCCCCTGGACT  
GATGAGTTTTGCTGTATCAACCTGTAAGGAGAAGCTCTCTCCGGATGGCTATGGGAATGAAAGAATCCGACTTCTA  
CTCTCACACAGCCACCGTGAAAGTCCTGGAGTAAATGTGCTGTGTACAGAAGAGAGAGAAGGAAGCAGGCTGGC  
ATGTTCACTGGGCTGGTGTACGACAGAGAACCTGACAGTCACTGGCCAGTTATCACTTCAGATTACAAATCACA  
CAGAGCATCTGCCTGTTTTCAATCACAAGAGAACAAAACCAAAATCTATAAAGATATTCTGAAATATGACAGAA  
TTTGACAAATAAAAGCATAAACGTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

217/615

**FIGURE 216**

MATWDEKAVTRRAKVAPAERMSKFLRHFTVVGGDYHAWNINYKKWENEEEEEEEEQPPPTPVSGEEGRAAAPDVA  
PAPGPAPRAPLDFRGMLRKLFSSSHREQVIIICLVVLDALLVLAELILDKIIQPDKNNYAAMVFHYMSITILVFF  
MMEIIFKLFVFRLLSFTTSLRSWMPVVVVVSFILDIVLLFQEHQFEALGLLLRLLRVARIINGIIISVKTRSE  
RQLRLKQMNVLAAKIQHLEFSCSEKPLD

218/615

**FIGURE 217**

GGAAGGCAGCGGCAGCTCCACTCAGCCAGTACCCAGATACGCTGGGAACCTTCCCCAGCCATGGCTTCCCTGGGG  
CAGATCCTCTTCTGGAGCATAATTAGCATCATCATTATTCTGGCTGGAGCAATTGCACTCATCATTGGCTTTGGT  
ATTTTCAGGGAGACACTCCATCACAGTCACTACTGTGCGCTCAGCTGGGAACATTGGGGAGGATGGAATCCTGAGC  
TGCATTTTTGAACCTGACATCAAACCTTTCTGATATCGTGATACAATGGCTGAAGGAAGGTGTTTTAGGCTTGGTC  
CATGAGTTCAAAGAAGGCAAAGATGAGCTGTGCGAGCAGGATGAAATGTTGAGAGGCCGGACAGCAGTGTGCT  
GATCAAGTGATAGTTGGCAATGCCTCTTTGCGGCTGAAAAACGTGCAACTCACAGATGCTGGCACCTACAAATGT  
TATATCATCACTTCTAAAGGCAAGGGGAATGCTAACCTTGAGTATAAACTGGAGCCTTCAGCATGCCGGAAGTG  
AATGTGGACTATAATGCCAGCTCAGAGACCTTGCGGTGTGAGGCTCCCCGATGGTTCCCCCAGCCCACAGTGGTC  
TGGGCATCCCAAGTTGACCAGGGAGCCAACCTTCTCGGAAGTCTCCAATACCAGCTTTGAGCTGAACCTCTGAGAAT  
GTGACCATGAAGGTTGTCTGTGTCTTACAATGTTACGATCAACAACACATACTCCTGTATGATTGAAAAATGAC  
ATTGCCAAAGCAACAGGGGATATCAAAGTGACAGAATCGGAGATCAAAGGCGGAGTCACCTACAGCTGCTAAAC  
TCAAAGGCTTCTCTGTGTGTCTCTTCTTTTCCATCAGCTGGGCACCTTCTGCCTCTCAGCCCTTACCTGATG  
CTAAAATAATGTGCCTTGGCCACAAAAAAGCATGCAAAGTCATTGTTACAACAGGGATCTACAGAACTATTTAC  
CACCAGATATGACCTAGTTTTATATTTCTGGGAGGAAATGAATTCATATCTAGAAGTCTGGAGTGAGCAAACAAG  
AGCAAGAAACAAAAAGAAGCCAAAAGCAGAAGGCTCCAATATGAACAAGATAAATCTATCTTCAAAGACATATTA  
GAAGTTGGGAAAATAATTTCATGTGAAGTAGACAAGTGTGTTAAGAGTGATAAGTAAATGCACGTGGAGACAAGT  
GCATCCCCAGATCTCAGGGACCTCCCCCTGCCTGTCACCTGGGGAGTGAGAGGACAGGATAGTGCATGTTCTTTG  
TCTCTGAATTTTTAGTTATATGTGCTGTAATGTTGCTCTGAGGAAGCCCCTGGAAAGTCTATCCCAACATATCCA  
CATCTTATATTCCACAAATTAAGCTGTAGTATGTACCCTAAGACGCTGCTAATTGACTGCCACTTCGCAACTCAG  
GGGCGGCTGCATTTTAGTAATGGGTCAAATGATTCACCTTTTTATGATGCTTCCAAAGGTGCCTTGGCTTCTCTTC  
CCAACTGACAAATGCCAAAGTTGAGAAAAATGATCATAATTTTAGCATAAACAGAGCAGTCGGGGACACCGATTT  
TATAAATAAAGTGAACCTTCTTTTTTAACAA  
AAAAAAA

219/615

**FIGURE 218**

MASLGQILFWSIIISIIIIILAGAIALIIGFGISGRHSITVTTVASAGNIGEDGILSCTFEPDIKLSDIVIQWLKEG  
VLGLVHEFKEGKDELSEQDEMFRGRTAVFADQVIVGNASLRLKNVQLTDAGTYKCYIITSKGKGNANLEYKTGAF  
SMPEVNVVDYNASSETLRCEAPRWFPQPTVVWASQVDQGANFSEVSNTSFELNSENVTMKVSVLYNVTINNTYSC  
MIENDIAKATGDIKVTSEIKRRSHLQLLNSKASLCVSSFFAISWALLPLSPYLMLK

FIGURE 219

GAATTTGTAGAAGACAGCGGCGTTGCCATGGCGGCGTCTCTGGGGCAGGTGTTGGCTCTGGTGTCTGGTGGCCGCTCTGTGGGGTGGCACGCAGCCGCTGCTGAAGCGGGCCTCCGCCGGCCTGCAGCGGGTTCATGAGCCGACCTGGGCCCAGCAGTTGCTACAGGAGATGAAGACCCTCTTCTTGAATACTGAGTACCTGATGCCCTTTCTCCTCAACCAGTGTGGATCCCTTCTCTATTACCTCACCTTGGCATCGACAGATCTGACCCTGGCTGTGCCCATCTGTAACCTCTCTGGCTATCATCTTACACATGATTGTTGGGAAGGCCCTTGAGAAGATATTTGGTGGAAACGTAAGTAGACTACTGCGAGTGGCGGACGCAGCTCTGTGGATCTCGACATACCTGTGTAGTTCTTCCCGAAGCCCATCTCCCGAGATGGGTGAGGACACGGCCCTTTTCCCATCTCGCCCTTTCCCTCTGCAGCTGTTTTGCTTCTTGTGGCCATCAGAGTTCCCTTCCTTGACAGTCTGGAGAAAGACAGAGGCTGGGGTTTGGGATTGAAGACCAGACCCCATCTGAGCCCTTCTCCA GCCCTGTACCAGCTCCTACTGGCATGGCTGAGCTCAGACCCTCCTGATTTCTGCCTATTATCCCAGGAGCAGTTGCTGGCATGGTGTCTACCGTGATAGGAATTTCACTCTGCATCACAAAGCTCAGTGAGTAAGACCCAGGGGCCAACAGTCTACCCTTTTGAAGTGGGCGCAACCCACTTCCAGCTCTGCTGCCTCCAGGAAGCCCTGGGCCATGAAGTGCTGGCA GTGAGCGGATGGACCTAGCACTTCCCCCTCTGCGCCTTAGCTTCTCTCTCTATGGGATAACAGCTACCTCA TGGATCACAAATAAGAGAACAAGAGTGAAAGAGTTTTGTAACCTTCAAGTGCTGTTTCAGCTGCGGGGATTTAGCAC AGGAGACTCTACGCTCACCCCTCAGCAACCTTTCTGCCCCAGCAGCTCTCTTCTGCTAACATCTCAGGCTCCAG CCCAGCCACCATTACTGTGGCCTGATCTGGACTATCATGGTGGCAGGTTCCATGGACTGCAGAACTCCAGCTGCA TGGAAAGGGCCAGCTGCAGACTTTGAGCCAGAAATGCAAACGGGAGGCCTCTGGGACTCAGTCAGAGCGCTTTGG CTGAATGAGGGGTGGAACCGAGGGAAGAAGGTGCGTCCGAGTGGCAGATGCAGGAAATGAGCTGTCTATTAGCCT TGCCTGCCCCACCCTAGGTAGGTAGGCAGAAATCCTCACTGCCAGCCCTCTTAAACAGGTAGAGAGCTGTGAGCCC CAGCCCCACTGACTCCAGCACACCTGGCGAGTAGTGTCTCAATAAATCTATGTAAACAGACAAAAA

221/615

**FIGURE 220**

MAASLGQVLALVLVAALWGGTQPLLKRASAGLQRVHEPTWAQQLQEMKTLFLNTEYLMFPLLNQCGSILLYLTL  
ASTDLTLAVPICNSLAIIFTLIVGKALGEDIGGKRKLDYCECGTQLCGSRHTCVSSFPEPISPEWVRTRPFPILP  
FPLQLFCFLVAIRVPFPWTVVRKTEAGVWD



222/615

**FIGURE 221**

CTTCTGTAGGACAGTCACCAGGCCAGATCCAGAAGCCTCTCTAGGCTCCAGCTTTCTCTGTGGAAGATGACAGCA  
ATTATAGCAGGACCCCTGCCAGGCTGTGCGAAAAGATTCCGCAATAAAACTTTGCCAGTGGGAAGTACCTAGTGAAA  
CGGCCTAAGATGCCACTTCTTCTCATGTCCCAGGCTTGAGGCCCTGTGGTCCCCATCCTTGGGAGAAGTCAGCTC  
CAGCACCATGAAGGGCATCCTCGTTGCTGGTATCACTGCAGTGCTTGTGTCAGCTGTAGAATCTCTGAGCTGCGT  
GCAGTGTAATTCATGGGAAAAATCCTGTGTCAACAGCATTGCCTCTGAATGTCCCTCACATGCCAACACCAGCTG  
TATCAGCTCCTCAGCCAGCTCCTCTCTAGAGACACCAGTCAGATTATACCAGAATATGTTCTGCTCAGCGGAGAA  
CTGCAGTGAGGAGACACACATTACAGCCTTCACTGTCCACGTGTCTGCTGAAGAACACTTTCATTTTGTAAAGCCA  
GTGCTGCCAAGGAAAGGAATGCAGCAACACCAGCGATGCCCTGGACCCTCCCCTGAAGAACGTGTCCAGCAACGC  
AGAGTGCCCTGCTTGTATGAATCTAATGGAATTCTGTGCTGGGAAGCCCTGGAAATGCTATGAAGAAGAACA  
GTGTGCTTTCTAGTTGCAGAACTTAAGAATGACATTGAGTCTAAGAGTCTCGTGCTGAAAGGCTGTTCCAACGT  
CAGTAACGCCACCTGTCACTTCCTGTCTGGTGAAAACAAGACTCTTGGAGGAGTCATCTTTCGAAAGTTTGAGTG  
TGCAAATGTAAACAGCTTAACCCCCACGTCTGCACCAACCACTTCCACAACGTGGGCTCCAAAGCTTCCCTCTA  
CCTCTTGGCCCTTGCCAGCCTCCTTCTTCGGGGACTGCTGCCCTTGAGGTCTCTGGGGCTGCACTTTGCCAGCACC  
CCATTTCTGCTTCTCTGAGGTCCAGAGCACCCCTGCGGTGCTGACACCCTCTTCCCTGCTCTGCCCCGTTTAA  
CTGCCCAGTAAGTGGGAGTCACAGGTCTCCAGGCAATGCCGACAGCTGCCTTGTTCTTCATTATTAAAGCACTGG  
TTCATTCAC TGCCAAAAAAAAAAAAAAAAAAAAAAAAA

223/615

**FIGURE 222**

MKGILVAGITAVLVAAVESLSCVQCNSWEKSCVNSIASECPSHANTSCISSSASSSLETPVRLYQNMFCSAENCS  
EETHITAFTVHVSAAEHFHFVSQCCQGKECSNTSDALDPPLKNVSSNAECPACYESNGTSCRGKPKWCYEEEQCV  
FLVAELKNDIESKSLVLKGCSNVSNATCQFLSGENKTLGGVIFRKFEKANVNSLTPTSAPTTSNHNVGSKASLYLL  
ALASLLLRGLLP

224/615

**FIGURE 223**

GGCCTCGGTTCAAACGACCCGGTGGGTCTACAGCGGAAGGGAGGGAGCGAAGGTAGGAGGCAGGGCTTGCCTCAC  
TGGCCACCCTCCCAACCCCAAGAGCCCAGCCCCATGGTCCCCGCCGCCGCCGCTGCTGTGGGTCCCTGCTGCTG  
AATCTGGGTCCCCGGGCGGCGGGGGCCCAAGGCCTGACCCAGACTCCGACCGAAATGCAGCGGGTCAGTTTACGC  
TTTGGGGGCCCCATGACCCGCAGCTACCGGAGCACCGCCCGACTGGTCTTCCCCGGAAGACAAGGATAATCCTA  
GAGGACGAGAATGATGCCATGGCCGACGCCGACCGCCTGGCTGGACCAGCGGCTGCCGAGCTCTTGGCCGCCACG  
GTGTCCACCGGCTTTAGCCGGTCGTCCGCCATTAACGAGGAGGATGGGTCTTCAGAAGAGGGGGTTGTGATTAAT  
GCCGGAAGGATAGCACCAGCAGAGAGCTTCCAGTGCGACTCCAATACAGCGGGGAGTTCCAGCACGAGGTTT  
ATAGCCAATAGTCAGGAGCCTGAAATCAGGCTGACTTCAAGCCTGCCGCGCTCCCCGGGAGGTCTACTGAGGAC  
CTGCCAGGCTCGCAGGCCACCCTGAGCCAGTGGTCCACACCTGGGTCTACCCCGAGCCGGTGGCCGTCACCTCA  
CCCACAGCCATGCCATCTCCTGAGGATCTGCGGCTGGTGCTGATGCCCTGGGGCCCGTGGCACTGCCACTGCAAG  
TCGGGCACCATGAGCCGAGCCGGTCTGGGAAGCTGCACGGCCTTTCGGGGCGCCTTCGAGTTGGGGCGCTGAGC  
CAGCTCCGCACGGAGCACAAAGCCTTGCACCTATCAACAATGTCCCTGCAACCGACTTCGGGAAGAGTGCCCCCTG  
GACACAAGTCTCTGTACTGACACCAACTGTGCCTCTCAGAGCACCACAGTACCAGGACCACCACTACCCCCCTC  
CCCACCATCCACCTCAGAAGCAGTCCCAGCCTGCCACCCGCCAGCCCCCTGCCAGCCCTGGCTTTTGGAAACGG  
GTCAGGATTGGCCTGGAGGATATTTGGAATAGCCTCTCTTCAGTGTTACAGAGATGCAACCAATAGACAGAAAC  
CAGAGGTAATGGCCACTTCATCCACATGAGGAGATGTCAGTATCTCAACCTCTCTTGCCCTTTCATCCTAGCAC  
CCACTAGATATTTTTAGTACAGAAAAACAAAACCTGGAAAAACACAA

225/615

**FIGURE 224**

MVPAAGALLWVLLLNLPRAAGAQGLTQTPTEMQRVSLRFGGPMTRSYRSTARTGLPRKTRIILEDENDAMADAD  
RLAGPAAAEELLAATVSTGFSSRSAINEDGSSEEGVVINAGKDSTSRELPSATPNTAGSSSTRFIANSQEPEIRL  
TSSLPRSPGRSTEDLPGSQATLSQWSTPGSTPSRWSPSPPTAMPSPEDLRLVLMFPGPWHCHCKSGTMSRSRSGK  
LHGLSGRLRVGALSQLRTEHKPCTYQQCPNRLREECPLDTSLCDTNCASQSTTSTRTTTTPFPTIHLRSSPSL  
PPASPCPALAFWKRVRIQLEDIWNLSLSSVFTEMQFIDRNQR

226/615

**FIGURE 225**

CCCGGGTCGACCCACGCGTCCGGGGAGAAAGGATGGCCGGCCTGGCGGGCGGGTTGGTCCTGCTAGCTGGGGCAG  
CGGCGCTGGCGAGCGGCTCCCAGGGCGACCGTGAGCCGGTGTAACGCGACTGCGTACTGCAGTGCGAAGAGCAGA  
ACTGCTCTGGGGGCGCTCTGAATCACTCCGCTCCCGCCAGCCAATCTACATGAGTCTAGCAGGCTGGACCTGTC  
GGGACGACTGTAAGTATGAGTGTATGTGGGTCAACGTTGGGCTCTACCTCCAGGAAGGTCACAAAGTGCCTCAGT  
TCCATGGCAAGTGGCCCTTCTCCCGGTTCTGTTCTTTCAAGAGCCGGCATCGGCCGTGGCCTCGTTTCTCAATG  
GCCTGGCCAGCCTGGTGATGCTCTGCCGCTACCGCACCTTCGTGCCAGCCTCCTCCCCATGTACCACACCTGTG  
TGGCCTTCGCCTGGGTGTCCCTCAATGCATGGTTCTGGTCCACAGTCTTCCACACCAGGGACACTGACCTCACAG  
AGAAAATGGACTACTTCTGTGCCTCCACTGTGCATCCTACACTCAATCTACCTGTGCTGCGTCAGGACCGTGGGGC  
TGCAGCACCCAGCTGTGGTCAGTGCCTTCGGGCTCTCCTGCTGCTCATGCTGACCGTGCACGTCTCCTACCTGA  
GCCTCATCCGCTTCGACTATGGCTACAACCTGGTGGCCAACGTGGCTATTGGCCTGGTCAACGTGGTGTGGTGGC  
TGGCCTGGTGCCTGTGGAAACAGCGGGCGGCTGCCTCACGTGCGCAAGTGGTGGTGGTCTTGTGCTGCAGG  
GGCTGTCCCTGCTCGAGCTGCTTGACTTCCCACCGCTCTTCTGGGTCTGGATGCCCATGCCATCTGGCACATCA  
GCACCATCCCTGTCCACGTCTCTTTTCAGCTTTCTGGAAGATGACAGCCTGTACCTGCTGAAGGAATCAGAGG  
ACAAGTTCAGCTGGACTGAAGACCTTGGAGCGAGTCTGCCCCAGTGGGGATCCTGCCCCCGCCCTGCTGGCCCTC  
CCTTCTCCCCTCAACCTTGAGATGATTTTCTCTTTTCAACTTCTTGAACCTTGGACATGAAGGATGTGGGCCAG  
AATCATGTGGCCAGCCACCCCTGTTGGCCCTCACCAGCCTTGGAGTCTGTTCTAGGGAAGGCCTCCCAGCATC  
TGGGACTCGAGAGTGGGCAGCCCTCTACCTCCTGGAGCTGAACTGGGGTGGAACTGAGTGTGTTCTTAGCTCTA  
CCGGGAGGACAGCTGCCTGTTTCTCCCCACCAGCCTCCTCCCCACATCCCCAGCTGCCTGGCTGGGTCTGAAG  
CCCTCTGTCTACCTGGGAGACCAGGGACCACAGGCCTTAGGGATACAGGGGGTCCCCTTCTGTTACCACCCCCA  
CCCTCCTCCAGGACACCACTAGGTGGTGTGGATGCTTGTCTTTGGCCAGCCAAGGTTACGGCGATTCCTCCCC  
ATGGGATCTTGAGGGACCAAGCTGCTGGGATTGGGAAGGAGTTTACCCTGACCGTTGCCCTAGCCAGGTTCCCA  
GGAGGCCTCACCATACTCCCTTTCAGGGCCAGGGCTCCAGCAAGCCCAGGGCAAGGATCCTGTGCTGTCTGTGG  
TTGAGAGCCTGCCACCGTGTGTGCGGAGTGTGGGCCAGGCTGAGTGCATAGGTGACAGGGCCGTGAGCATGGGCC  
TGGGTGTGTGTGAGCTCAGGCCTAGGTGCGCAGTGTGGAGACGGGTGTGTGCGGGGAAGAGGTGTGGCTTCAAAG  
TGTGTGTGTGCAGGGGGTGGGTGTGTTAGCGTGGGTTAGGGGAACGTGTGTGCGCGTGTGGTGGGCATGTGAGA  
TGAGTGACTGCCGGTGAATGTGTCCACAGTTGAGAGGTTGGAGCAGGATGAGGGAATCCTGTACCATCAATAAT  
CACTTGTGGAGCGCCAGCTCTGCCAAGACGCCACCTGGGCGGACAGCCAGGAGCTCTCCATGGCCAGGCTGCCT  
GTGTGCATGTTCCCTGTCTGTTGCCCTTTGCCCGCCTCCTGCAAACCTCACAGGGTCCCCACACAACAGTGCCC  
TCCAGAAGCAGCCCCTCGGAGGCAGAGGAAGGAAATGGGGATGGCTGGGGCTCTCTCCATCCTCCTTTTCTCCT  
TGCCTTCGCATGGCTGGCCTTCCCCTCCAAAACCTCCATTCCCCTGCTGCCAGCCCCCTTGGCCATAGCCTGATTT  
TGGGGAGGAGGAAGGGGCGATTTGAGGGAGAAGGGGAGAAAGCTTATGGCTGGGTCTGGTTTCTTCCCTTCCCAG  
AGGGTCTTACTGTTCCAGGGTGGCCCCAGGGCAGGCAGGGGCCACACTATGCCTGTGCCCTGGTAAAGGTGACCC  
CTGCCATTTACCAGCAGCCCTGGCATGTTCTGCCCCACAGGAATAGAATGGAGGGAGCTCCAGAACTTTCCAT  
CCCAAAGGCAGTCTCCGTGGTTGAAGCAGACTGGATTTTTGCTCTGCCCTGACCCCTTGTCCCTCTTTGAGGGA  
GGGGAGCTATGCTAGGACTCCAACCTCAGGGACTCGGGTGGCCTGCGCTAGCTTCTTTTGATACTGAAAACCTTTT  
AAGGTGGGAGGGTGGCAAGGGATGTGCTTAATAAATCAATTCCAAGCCTCAAAAAAAAAAAAAAAAAA

227/615

**FIGURE 226**

MAGLAARLVLLAGAAALASGSQGDREPVYRDCVLQCEEQNCSGGALNHFRSRQPIYMSLAGWTCRDDCKYECMWV  
TVGLYLQEGHKVPQFHGKWPFSRFLFFQEPASAVASFNLGLASLVMLCRYRTFVPASSPMYHTCVAFWVSLNAW  
FWSTVFHTRDLDLTKMDYFCASTVILHSIYLCCVRTVGLQHAPVVSFAFRALLLMLTVHVSYSLSLIRFDYGYNL  
VANVAIGLVNVVWVWLAWCLWNQRRLPVVRKCVVVVLLQLGLSLELLDFFPLFWVLDAAHAIWHISTIPVHVLFSS  
FLEDDSLYLLKESEDKFKLD

**Important features:****Signal peptide:**

amino acids 1-20

**Transmembrane domains:**

amino acids 105-123, 138-156, 169-185, 193-209, 221-240, 256-272

**N-glycosylation site.**

amino acids 40-44

**N-myristoylation site.**

amino acids 43-49

**CUB domain proteins profile.**

amino acids 285-302

**Amiloride-sensitive sodium channels proteins.**

amino acids 162-186

228/615

**FIGURE 227**

TTCGGCTTCCGTAGAGGAAGTGGCGCGGACCTTCATTTGGGGTTTCGGTTCCCCCCTTCCCCTTCCCCGGGGTC  
TGGGGGTGACATTGCACCGCGCCCCCTCGTGGGGTCGCGTTGCCACCCACGCGGACTCCCCAGCTGGCGCGCCCC  
TCCCATTTCCTGTCTCTGGTCAGGCCCCACCCCCCTTCCACCTGACCAGCCATGGGGGCTGCGGTGTTTTTCG  
GCTGCACTTTTCGTGCGGTTTCGGCCCCGGCCTTCGCGCTTTTCTTGATCACTGTGGCTGGGGACCCGCTTCGCGTTA  
TCATCCTGGTCGAGGGGCATTTTTCTGGCTGGTCTCCCTGCTCCTGGCCTCTGTGGTCTGGTTCATCTTGGTCC  
ATGTGACCGACCGGTGAGATGCCCCGGCTCCAGTACGGCCTCCTGATTTTTGGTGCTGCTGTCTCTGTCTCTAC  
AGGAGGTGTTCCGCTTTGCTACTACAAGCTGCTTAAGAAGGCAGATGAAGGGTTAGCATCGCTGAGTGAGGACG  
GAAGATCACCCATCTCCATCCGCCAGATGGCCTATGTTTCTGGTCTCTCCTTCGGTATCATCAGTGGTGTCTTCT  
CTGTTATCAATATTTTGGCTGATGCACTTGGGCCAGGTGTGGTTGGGATCCATGGAGACTCACCTATTACTTCC  
TGACTTCAGCCTTTCTGACAGCAGCCATTATCCTGCTCCATACCTTTTGGGGAGTTGTGTTCTTTGATGCCTGTG  
AGAGGAGACGCTACTGGGCTTTGGGCCTGGTGGTTGGGAGTCACCTACTGACATCGGGACTGACATTCTGAACC  
CCTGGTATGAGGCCAGCCTGCTGCCCATCTATGCAGTCAGTGTTCCTGCGGCTCTGGGCCTTCATCACAGCTG  
GAGGGTCCCTCCGAAGTATTCAGCGCAGCCTCTTGTGTAAGGACTGACTACCTGGACTGATCGCCTGACAGATCC  
CACCTGCCTGTCCACTGCCCATGACTGAGCCCCAGCCCCAGCCCGGTCCATTGCCCACATTCTCTGTCTCCTTCT  
CGTCGGTCTACCCCACTACCTCCAGGGTTTTGCTTTGTCTTTTGTGACCGTTAGTCTCTAAGCTTTACAGGAG  
CAGCCTGGGTTGAGCCAGTCAGTGACTGGTGGGTTTGAATCTGCACTTATCCCCACCACCTGGGGACCCCTTGT  
TGTGTCCAGGACTCCCCCTGTGTGAGTGCTCTGCTCTCACCTGCCCCAAGACTCACCTCCCTTCCCCTCTGCAGG  
CCGACGGCAGGAGGACAGTCGGGTGATGGTGTATTCTGCCCTGCGCATCCCACCCGAGGACTGAGGGAACCTAGG  
GGGGACCCCTGGGCCTGGGGTGGCCTCCTGATGTCTCGCCCTGTATTTCTCCATCTCCAGTTCTGGACAGTGCA  
GGTTGCCAAGAAAAGGGACCTAGTTTAGCCATTGCCCTGGAGATGAAATTAATGGAGGCTCAAGGATAGATGAGC  
TCTGAGTTTCTCAGTACTCCCTCAAGACTGGACATCTTGGTCTTTTCTCAGGCCTGAGGGGGAACCATTTTTGG  
AACCTCCTTGGGCTATATTTTTCTCTCCTCGAGTTGCTCCTCATGGCTGGGCTCATTTTCGGTCCCTTTCTCCTTGG  
TCCCAGACCTTGGGGGAAGGAAGGAAGTGCATGTTTGGGAACTGGCATTACTGGAATAATGGTTTTAACCTCC  
TTAACACCAGCATCCCTCCTCTCCCCAAGGTGAAGTGGAGGGTGCTGTGGTGAGCTGGCCACTCCAGAGCTGCA  
GTGCCACTGGAGGAGTCAGACTACCATGACATCGTAGGGAAGGAGGGGAGATTTTTTTGTAGTTTTTAATTGGGG  
TGTTGGGAGGGGCGGGGAGGTTTTCTATAAACTGTATCATTTTCTGCTGAGGGTGGAGTGTCCCATCCTTTTAATC  
AAGGTGATTGTGATTTTGAATAATAAAGAATTTGTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

229/615

**FIGURE 228**

MGAAVFFGCTFVAFGPAFALFLITVAGDPLRVIIILVAGAFFWLVSILLASVVWFILVHVTDRSDARLQYGLLIFG  
AAVSVLLQEVFRFAYYKLLKKADEGLASLSEDRSPISIRQMAYVSGLSFGIISGVFSVINILADALGPGVVGIIH  
GDSPYYFLTSAFLTAAIILLHTFWGVVFFDACERRRYWALGLVVGSHLLTSGLTFLNPWYEASLLPIYAVTVSMG  
LWAFITAGGSLRSIQRSLLCKD



230/615

**FIGURE 229**

CGGGAGGCTGGGTGCTCATGATCCGGACCCATTGTGCGGCTCTGCCCATCGCCTGCTCCTCCCAGGCTCCCGCG  
GCCGACCCCCGCGCAAC**ATG**CAGCCACGGGCGCGAGGGTTCCCGCGCGCTCAGCCGGCGGTATCTGCGGCGTC  
TGCTGCTCCTGCTACTGCTGCTGCTGCTGCGGCAGCCCGTAACCCGCGCGGAGACCACGCCGGGCGCCCCAGAG  
CCCTCTCCACGCTGGGCTCCCCAGCCTCTTACACAGCCGGGTGTCCCAGCGCCCTCACTACCCAGGCCTCA  
CTACGCCAGGCACCCCCAAAACCCTGGACCTTCGGGGTCGCGCGCAGGCCCTGATGCGGAGTTTCCCACTCGTGG  
ACGGCCACAATGACCTGCCCCAGGTCCTGAGACAGCGTTACAAGAATGTGCTTCAGGATGTTAACCTGCGAAATT  
TCAGCCATGGTCAGACCAGCCTGGACAGGCTTAGAGACGGCCTCGTGGGTGCCAGTTCTGGTCAGCCTCCGTCT  
CATGCCAGTCCCAGGACCAGACTGCCGTGCGCCTCGCCCTGGAGCAGATTGACCTCATTACCCGCATGTGTGCCT  
CCTACTCTGAACTCGAGCTTGTGACCTCAGCTGAAGGTCTGAACAGCTCTCAAAAGCTGGCCTGCCTCATTGGCG  
TGNAGGGTGGTCACTCACTGGACAGCAGCCTCTCTGTGCTGCGCAGTTTCTATGTGCTGGGGGTGCGCTACCTGA  
CACTTACCTTCACCTGCAGTACACCATGGGCAGAGAGTTCCACCAAGTTCAGACACCACATGTACACCAACGTCA  
GCGGATTGACAAGCTTTGGTGAGAAAGTAGTAGAGGAGTTGAACCGCCTGGGCATGATGATAGATTTGTCTTATG  
CATCGGACACCTTGATAAGAAGGGTCCTGGAAGTGTCTCAGGCTCCTGTGATCTTCTCCCACTCAGCTGCCAGAG  
CTGTGTGTGACAATTTGTTGAATGTTCCCGATGATATCCTGCAGCTTCTGAAGAACGGTGGCATCGTGATGGTGA  
CACTGTCCATGGGGGTGCTGCAGTGCAACCTGCTTGCTAACGTGTCCACTGTGGCAGATCACTTTGACCACATCA  
GGGCAGTCATTGGATCTGAGTTCATCGGCATTGGTGGAAATTATGACGGGACTGGCCGGTTCCCTCAGGGGCTGG  
AGGATGTGTCCACATACCCAGTCCTGATAGAGGAGTTGCTGAGTCGTASCTGGAGCGAGGAAGAGCTTCAAGGTG  
TCCTTCGTGGAACCTGCTGCGGGTCTTCAGACAAGTGGAAGAGGTGAGAGAGGAGAGCAGGGCGCAGAGCCCCG  
TGGAGGCTGAGTTTCCATATGGGCAACTGAGCACATCCTGCCACTCCACCTCGTGCCTCAGAATGGACACCAGG  
CTACTCATCTGGAGGTGACCAAGCAGCCAACCAATCGGGTCCCCTGGAGGTCTCAAATGCCTCCCCATACCTTG  
TTCCAGGCCTTGTGGCTGCTGCCACCATCCCAACCTTCACCCAGTGGCTCTGCT**TGA**CACAGTCGGTCCCCGCAGA  
GGTCACTGTGGCAAAGCCTCACAAAGCCCCCTCTCCTAGTTCATTACAAAGCATATGCTGAGAATAAACATGTTA  
CACATGGAAAA

231/615

**FIGURE 230**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59817
><subunit 1 of 1, 487 aa, 1 stop, 2 unknown
><MW: 53569.32, pI: 7.68, NX(S/T): 5
MQPTGREGSRALSRRYLRRLLLLLLLLLLLLLRQPVTRAETTPGAPRALSTLGSPSLFTTPGVPSALTTPGLTTPGTP
KTLDLRGRAQALMRSFPLVDGHNDLPQVLRQRYKNVLQDVNLRNFSGQTSLDRLRDGLVGAQFWSASVSCQSQD
QTAVRLALEQIDLIHRMCASYSELELVTSAEGLNSSQKLACLIGVXGGHSLDSSLSVLRSFYVLGVRYLTLTFTC
STPWAESSTKFRHHMYTNVSGLTSGEKKVVEELNRLGMMIDLSYASDTLIRRVLEVSQAPVIFSHSAARAVCDNL
LNVPPDILQLLKNGGIVMVTLSMGVLQCNLLANVSTVADHFDHIRAVIGSEFIGIGGNYDGTGRFPQGLEDVSTY
PVLIEELLSRXWSEELQGVLRGNLLRVFRQVEKVRRESRAQSPVEAEFPYQQLSTSCHSHLVPQNGHQATHLEV
TKQPTNRVPWRSSNASPYLVPGLVAAATIPTFTQWLC
```

**Important features of the protein:****Signal peptide:**

amino acids 1-36

**Transmembrane domain:**

amino acids 313-331

**N-glycosylation sites.**

amino acids 119-122, 184-187, 243-246 and 333-336

**N-myristoylation sites.**

amino acids 41-46, 59-64, 73-78, 133-138, 182-187, 194-199, 324-329, 354-359, 357-362, 394-399, 427-432 and 472-477.

**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 136-146

232/615

**FIGURE 231**

GCTCTGGCCGGCCCCGGCGATTGGTCACCGCCCGCTAGGGGACAGCCCTGGCCTCCTCTGATTGGCAAGCGCTGG  
CCACCTCCCCACACCCCTTGCGAACGCTCCCCTAGTGGAGAAAAGGAGTAGCTATTAGCCAATTTCGGCAGGGCCC  
GCTTTTTAGAAAGCTTGATTTCCTTTGAAGATGAAAGACTAGCGGAAGCTCTGCCTCTTTCCCCAGTGGGCGAGGG  
AACTCGGGGCGATTGGCTGGGAAGTGTATCCACCCAAATGTACCGATTCTTCCTATGCAGGAAATGAGCAGAC  
CCATCAATAAGAAATTTCTCAGCCTGGCCGAAAATGGTTGGCCCCACGAAGCCACGACAAGTGGAGGCAAAGAGG  
GTTGCTCAACGCCCCGCTCATTGGAAAACCAAATCAGATCTGGGACCTATATAGCGTGGCGGAGGCGGGGCGAT  
GATTGTCGCGCTCGCACCCACTGCAGCTGCGCACAGTCGCATTTCTTTCCCCGCCCCCTGAGACCCTGCAGACCA  
TCTGTCAATGGCGGCTGGGCTGTTTGGTTTGAGCGCTCGCCGTCTTTTGGCGGCAGCGGCGACGCGAGGGCTCCCG  
GCGCCCCGCTCCGCTGGGAATCTAGCTTCTCCAGGACTGTGGTCGCCCCGTCCGCTGTGGCGGAAAGCGGCCC  
CCAGAACCGACCACACCGTGGCAAGAGGACCCAGAACCCGAGGACGAAAAGTGTATGAGAAGAACCCAGACTCC  
CATGGTTATGACAAGGACCCCGTTTGGACGTCTGGAACATGCGACTTGTCTTCTTTGGCGCTCTCCATCATC  
CTGGTCCTTGGCAGCACCTTTGTGGCCTATCTGCCTGACTACAGGATGAAAGAGTGGTCCCGCCGCGAAGCTGAG  
AGGCTTGTGAAATACCGAGAGGCCAATGGCCTTCCCATCATGGAATCCAAGTCTTCGACCCAGCAAGATCCAG  
CTGCCAGAGGATGAGTGACCAGTTGCTAAGTGGGGCTCAAGAAGCACCGCCTTCCCCACCCCTGCCTGCCATTCT  
TGACCTCTTCTCAGAGCACCTAATTAAAGGGGCTGAAAGTCTGAA

233/615

**FIGURE 232**

MAAGLFGLSARRLLAAAATRGLPAARVRWESSFSRTVVAPSAVAGKRPPEFTTPWQEDPEPEDENLYEKNPDSHG  
YKDPVLDVWNMLVFFFGVSIILVLGSTFVAYLPDYRMKEWSRREAERLVKYREANGLPIMESNCFDPSKIQLPEDE

234/615

**FIGURE 233**

GCGGGGGCTATGCCGCTTGCTCTGCTCGTCCTGTTGCTCCTGGGGCCCGGCGGCTGGTGCCTTGACAGAACCCCCA  
CGCGACAGCCTGCGGGAGGAACTTGTATCACCCTCGCTGCCCTCCGGGGACGTAGCCGCCACATTCAGTTCCGC  
ACGCGCTGGGATTTCGGAGCTTCAGCGGGAAGGAGTGTCCCATTCAGGCTCTTTCCCAAAGCCCTGGGGCAGCTG  
ATCTCCAAGTATTCTCTACGGGAGCTGCACCTGTCTTACACAAAGGCTTTGGAGGACCCGATACTGGGGGCCA  
CCCTTCCTGCAGGCCCCATCAGGTGCAGAGCTGTGGGTCTGGTTCGAAGACACTGTCACTGATGTGGATAAATCT  
TGGAAGGAGCTCAGTAATGTCTCTCAGGGATCTTCTGCGCCTCTCTCAACTTCATCGACTCCACCAACACAGTC  
ACTCCCACTGCCTCCTTCAAACCCCTGGGTCTGGCCAATGACACTGACCACTACTTTCTGCGCTATGCTGTGCTG  
CCGCGGGAGGTGGTCTGCACCGAAAACCTCACCCCTGGAAGAAGCTCTTGGCCTGTAGTTCCAAGGCAGGCCTC  
TCTGTGCTGCTGAAGGCAGATCGCTTGTCCACACCAGCTACCCTCCAGGCAGTGCATATCCGCCCTGTTTGC  
AGAAATGCACGCTGTACTAGCATCTCCTGGGAGCTGAGGCAGACCCTGTCACTTGTATTTGATGCCTTCATCAG  
GGGCAGGGAAGAAAGACTGGTCCCTCTCCGGATGTTCTCCGAACCTCACGGAGCCCTGCCCCCTGGCTTCA  
GAGAGCCGAGTCTATGTGGACATCACCACTACAACCAGGACAACGAGACATTAGAGGTGCACCCACCCCGACC  
ACTACATATCAGGACGTCTCCTAGGCACTCGGAAGACCTATGCCATCTATGACTTGCTTGACACCGCCATGATC  
AACAACCTCTCGAAACCTCAACATCCAGCTCAAGTGGAAGAGACCCCCAGAGAATGAGGCCCCCAGTGCCCTTC  
CTGCATGCCCAGCGGTACGTGAGTGGCTATGGGCTGCAGAAGGGGGAGCTGAGCACACTGCTGTACAACACCCAC  
CCATACCGGGCCTTCCCGGTGCTGCTGCTGGACACCGTACCCTGGTATCTGCGGCTGTATGTGCACACCCCTCACC  
ATCACCTCCAAGGGCAAGGAGAACAACCAAGTTACATCCACTACCAGCCTGCCAGGACCGGCTGCAACCCAC  
CTCCTGGAGATGCTGATTAGCTGCCGGCCAACTCAGTCACCAAGGTTTCCATCCAGTTTGAGCGGGCGCTGCTG  
AAGTGGACCGAGTACACGCCAGATCCTAACCATGGCTTCTATGTAGCCCATCTGTCTCAGCGCCCTTGTGCCC  
AGCATGGTAGCAGCCAAGCCAGTGGACTGGGAAGAGAGTCCCCTCTTCAACAGCCTGTTCCAGTCTCTGATGGC  
TCTAACTACTTTGTGCGGCTCTACACGGAGCCGCTGCTGGTGAACCTGCCGACACCGGACTTCAGCATGCCCTAC  
AACGTGATCTGCCTCACGTGCACTGTGGTGGCCGTGTGCTACGGCTCCTTCTACAATCTCCTCACCCGAACCTTC  
CACATCGAGGAGCCCCGCACAGGTGGCCTGGCCAAGCGGCTGGCCAACCTTATCCGGCGCGCCGAGGTGTCCCC  
CCACTCTGATTCTTGGCCCTTTCAGCAGCTGCAGCTGCCGTTTCTCTCTGGGGAGGGGAGCCCAAGGGCTGTTTC  
TGCCACTTGCTCTCCTCAGAGTTGGCTTTTGAACCAAGTGCCCTGGACCAGGTGAGGGCTACAGCTGTGTTGT  
CCAGTACAGGAGCCACGAGCCAAATGTGGCATTGTGAATTTGAATTAACCTAGAAATTCATTTCTCACCTGTAGT  
GGCCACCTCTATATTGAGGTGCTCAATAAGCAAAAGTGGTGGTGGCTGCTGTATTGGACAGCACAGAAAAAGAT  
TTCCATCACACAGAAAGGTGCGCTGGCAGCACTGGCCAAGGTGATGGGGTGTGCTACACAGTGTATGTCACTGT  
GTAGTGGATGGAGTTTACTGTTTGTGGAATAAAAACGGCTGTTTCCGTGGAACCAAAAAA

235/615

**FIGURE 234**

MPLALLVLLLLGPGGWCLAEPPRDSLREELVITPLPSGDVAATFQFRTRWDSELQREGVSHYRLFPPKALGQLISK  
YSLRELHLSFTQGEWRTRYWGPPFLQAPSGAELWVWFQDTVTDVDSWKELSNVLSGIFCASLNFIDSTNTVTPT  
ASFKPLGLANDTDHYFLRYAVLPREVVCTENLTPWKLLPCSSKAGLSVLLKADRLFHTSYHSQAVHIRPVCRNA  
RCTSISWELRQTLSSVVFDAFITGQKKDWSLFRMFSTRTEPCPLASESRVYVDITTYNQDNETLEVHPPPTTTY  
QDVILGTRKTYAIYDLLDTAMINNSRNINIQLKWKRPENEAPPVPFLHAQRYVSGYGLQKELSTLLYNTHPYR  
AFPVLLLDTPWYLRLYVHTLTITSKGKENKPSYIHYQPAQDRLQPHLLEMLIQLPANSVTKVSIQFERALLKWT  
EYTPDPNHGFYVSPSVLSALVPSMVAAPVDWEESPLFNSLFPVSDGSNYFVRLYTEPLLVLNLPDPFSMPYNVI  
CLTCTVVAVCYGSFYNLLTRTFHIEEPRTGGLAKRLANLIRRARGVPPL

236/615

**FIGURE 235**

TGACGTCAGAATCACCATGGCCAGCTATCCTTACCGGCAGGGCTGCCAGGAGCTGCAGGACAAGCACCAGGAGC  
CCCTCCGGGTAGCTACTACCCTGGACCCCCCAATAGTGGAGGGCAGTATGGTAGTGGGCTACCCCTGGTGGTGG  
TTATGGGGGTCTGCCCCGAGGGCCTTATGGACCACCAGCTGGTGGAGGGCCCTATGGACACCCCAATCCTGG  
GATGTTCCCTCTGGAACTCCAGGAGGACCATATGGCGGTGCAGCTCCCGGGGGCCCCCTATGGTCAGCCACCTCC  
AAGTTCCTACGGTGCCCAGCAGCCTGGGCTTTATGGACAGGGTGGCGCCCCCTCCCAATGTGGATCCTGAGGCCTA  
CTCCTGGTTCCAGTCGGTGGACTCAGATCACAGTGGCTATATCTCCATGAAGGAGCTAAAGCAGGCCCTGGTCAA  
CTGCAATTGGTCTTCATTCAATGATGAGACCTGCCTCATGATGATAAACATGTTTGACAAGACCAAGTCAGGCCG  
CATCGATGTCTACGGCTTCTCAGCCCTGTGGAAATTCATCCAGCAGTGGGAAGAACCTCTCCAGCAGTATGACCG  
GGACCGCTCGGGCTCCATTAGCTACACAGAGCTGCAGCAAGCTCTGTCCCAATGGGCTACAACCTGAGCCCCA  
GTTACCCAGCTTCTGGTCTCCCGCTACTGCCACGCTCTGCCAATCCTGCCATGCAGCTTGACCGCTTCATCCA  
GGTGTGCACCCAGCTGCAGGTGCTGACAGAGGCCCTCCGGGAGAAGGACACAGCTGTACAAGGCAACATCCGGCT  
CAGCTTCGAGGACTTCGTCACCATGACAGCTTCTCGGATGCTATGACCCAACCATCTGTGGAGAGTGGAGTGCAC  
CAGGGACCTTTCTGGCTTCTTAGAGTGAGAGAAGTATGTGGACATCTCTTCTTTCTGTCCCTCTAGAAGAAC  
ATTCTCCCTTGCTTGATGCAACACTGTTCCAAAAGAGGGTGGAGAGTCTGCATCATAGCCACCAATAGTGAGG  
ACCGGGGCTGAGGCCACACAGATAGGGGCCTGATGGAGGAGAGGATAGAAGTTGAATGTCCTGATGGCCATGAGC  
AGTTGAGTGGCACAGCCTGGCACAGGAGCAGGTCCTTGTAAATGGAGTTAGTGTCCAGTCAGCTGAGCTCCACCC  
TGATGCCAGTGGTGAGTGTTTCATCGGCCTGTTACCGTTAGTACCTGTGTTCCCTCACCAGGCCATCCTGTCAAAC  
GAGCCCATTTTCTCCAAAGTGAATCTGACCAAGCATGAGAGAGATCTGTCTATGGGACCAGTGGCTTGGATTCT  
GCCACACCCATAAATCCTTGTGTGTTAACTTCTAGCTGCCTGGGGCTGGCCCTGCTCAGACAAATCTGCTCCCTG  
GGCATCTTTGGCCAGGCTTCTGCCCCCTGCAGCTGGGACCCCTCACTTGCCTGCCATGCTCTGCTCGGCTTCAGT  
CTCCAGGAGACAGTGGTCACCTCTCCCTGCCAATACTTTTTTAAATTTGCATTTTTTTTCATTTGGGGCCAAAAG  
TCCAGTGAAATTGTAAGCTTCAATAAAAGGATGAAACTCTGA

237/615

**FIGURE 236**

MASYPYRQGCPGAAGQAPGAPPGSYYPGPPNSGGQYGSGLPPGGGYGGPAPGGPYGPPAGGGPYGHPNPGMFPSG  
TPGGPYGGAAPGGPYGQPPSSYGAQQPGLYGQGGAPPNVDPEAYSWFQSVDSHSGYISMKELKQALVNCNWSS  
FNDETCLMMINMFDKTKSGRIDVYGFSA LWKFIQWKNLFQQYDRDRSGSISYTELQQALSQMGYNLSPQFTQLL  
VSRYCPRSANPAMQLDRFIQVCTQLQVLTEAFREKDTAVQGNIRLSFEDFVTMTASRML

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**N-glycosylation site.**

amino acids 147-150

**Casein kinase II phosphorylation sites.**

amino acids 135-138, 150-153, 202-205, 271-274

**N-myristoylation sites.**

amino acids 9-14, 15-20, 19-24, 33-38, 34-39, 39-44, 43-48, 61-66, 70-75, 78-83, 83-88, 87-92, 110-115



**FIGURE 237**

[illegible]

239/615

**FIGURE 238**

MQGRVAGSCAPLG LLLVCLHLPGLFARSIGVVEEKVSQNF GTNLPQLGQPSTGSPNSEHPQPALDPRSNDLARV  
PLKLSVPPSDGFPFAGGSAVQRWPPSWGLPAMDSWPPEDPWQMMAAAAE DRLGEALPEELSYLSSAAALAPGSGP  
LPGESSPDATGLSPEASLLHQDSESRRLPRSNSLGAGGKILSQRPPWSLIHRVLPDHPWGT LNPSVSWGGGGPGT  
GWGTRPMPHPEGIWGINNQPPGT SWGNINRYPGGSWGNINRYPGGSWGNINRYPGGSWGNIHLYPGINNPFP PGV  
LRPPGSSWNIPAGFPNPPSPRLQWG

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**Casein kinase II phosphorylation sites.**

amino acids 56-59, 155-158

**N-myristoylation sites.**amino acids 48-53, 220-225, 221-226, 224-229, 247-252, 258-263, 259-264, 269-  
274, 270-275, 280-285, 281-286, 305-310

240/615

**FIGURE 239**

GGGCGTCTCCGGCTGCTCCTATTGAGCTGTCTGCTCGCTGTGCCGCTGTGCCTGCTGTGCCGCGCTGTGCGCG  
CTGCTACCGCGTCTGCTGGACGCGGGAGACGCCAGCGAGCTGGTGATTGGAGCCCTGCCGAGAGCTCAAGCGCCC  
AGCTCTGCCCCAGGAGCCCAGGCTGCCCCGTGAGTCCCATAGTTGCTGCAGGAGTGGAGCCATGAGCTGCGTCCT  
GGGTGGTGTATCCCCCTTGGGGCTGCTGTTCTTGGTCTGCGGATCCCAAGGCTACCTCCTGCCCAACGTCACTCT  
CTTAGAGGAGCTGCTCAGCAAATACCAGCACAAACGAGTCTCACTCCCGGGTCCGAGAGCCATCCCCAGGGAGGA  
CAAGGAGGAGATCCTCATGCTGCACAACAAGCTTCGGGGCCAGGTGCAGCCTCAGGCCTCCAACATGGAGTACAT  
GGTGAGCGCCGGCTCCGGCCGAGAGGCTGGCACCGGGGTGGGGCCTGGGCCACCAGCCTGCTCTGTTCCCCAG  
CCAGCTCTGTTCCCCAGCCAGTGCCTGTGATGGCTGGCTCAGGGTCTCCTCTGGCAGGGGAGGATCCCCGGCTCTG  
TTCTGTTTTGTTTGTGTTGTTTTGAGACAGGGTCTCACTCTGCCACTGACGCTGGAGTGCAATGGCACAATCGTCA  
TGCCCTGAAACCTTAGACTCCCGGGGTAAAGCGATCCTGCTTCAGCCTCCCAAGTAGCTGGAACACAGGCATGC  
ACCATGGTGCCAGCTAGATTTTAAATATTTTGTGGAGATGGGGGTCTTGCTACGTTGCCAGGCTGGTCTTGAA  
CTCCTAGGCTCAAGCAATCCTCCTGCCTCAGCCTCTCAAAGTGCTAGGATTATAGGCATGAGTCACCCTGTCTGG  
CTCTGGCTCTGTTCTTAACATTCTGCCAAAACAACACACGTGGGTTCCTGTGCAGAGCCTGCCTCGTTGCCTTC  
ATGTCACTCTTGGTAGCTCCACTGGGAACACAGCTCTCAGCCTTTCCACCTGGAGGCAGAGTGGGGAGGGGGCC  
AGGGCTGGGCTTTGCTGATGCTGATCTCAGCTGTGCCACACGCTAGCTGCACCACCCTGACTTCTCCTTAGCCCC  
TGTGAGCCTCACTTTCCACTTGGAGAGTCCTTCCTCGGTGGTTGCCATGACTGTGAGATAAGTCGAGGCTGTGA  
AGGGCCCGGCACAGACTGACCTGCCTCCCCAACCCCTAGGCTTTGCTAACCGGGAAAGGAGCTAACGGTGACAGA  
AGACAGCCAAGGTCAACCCTCCCGGGTGATTGTGATGGGTGTTCCAGGTGTGGTTGGGCGATGCTGCTACTTGAC  
CCCAAGCTCCAGTGTGGAACCTTCCTTCCTGGCTGGTTTTCCAGAACTACAGAGGAATGGACCACAGTCTTCCAG  
GGTCCCTCCTCGTCCACCAACCGGGAGCCTCCACCTTGGCCATCCGTCAGCTATGAATGGCTTTTTAAACAAACC  
CACGTCCCAGCCTGGGTAACATGGTAAAGCCCCGTCTCTACAAAAAATCCAAGTTAGCCGGGCATGGTGGTGCG  
CACCTGTAGTCCCAGCTGCAGTGGGACTGAGGTGGAGGTGGAGGTGGGGGTGGGAGCTGAGGAAGGAGGATCGC  
TTGAGCCTGGGAAGTCGAGGCTGCAGTGAGCTGAGATTGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGAC  
CCTGTCTCAAAAA

241/615

## **FIGURE 240**

MSCVLGGVIPLGLLFLVCGSQGYLLPNVTLLLEELLSKYQHNEHSRVRRAIPREDKEEILMLHNKLRGQVQPQAS  
NMEYMVSAAGSGRRGWHRGWGLGHQPALFPSQLCSPASACDGWLRVSSGRGGSRLCSVLFVCFETGSHSATDAGVQ  
WHNRHALKP

**Important features:**

**Signal peptide:**

amino acids 1-22

**N-glycosylation site.**

amino acids 27-31, 41-45

**N-myristoylation site.**

amino acids 126-132, 140-146

**Amidation site.**

amino acids 85-89

242/615

**FIGURE 241**

AAGGAGAGGCCACCGGGACTTCAGTGTCTCCTCCATCCCAGGAGCGCAGTGGCCACTATGGGGTCTGGGCTGCCC  
CTTGTCCTCCTCTTGACCCCTCCTTGGCAGCTCACATGGAACAGGGCCGGGTATGACTTTGCAACTGAAGCTGAAG  
GAGTCTTTTCTGACAAATTCCTCCTATGAGTCCAGCTTCCTGGAATTGCTTGAAAAGCTCTGCCTCCTCCTCCAT  
CTCCCTTCAGGGACCAGCGTCACCCTCCACCATGCAAGATCTCAACACCATGTTGTCTGCAACACATTGACAGCCA  
TTGAAGCCTGTGTCCTTCTTGGCCCGGGCTTTTGGGCCGGGGATGCAGGAGGCAGGCCCCGACCCTGTCTTTCAG  
CAGGCCCCCACCTCCTGAGTGGCAATAAATAAAATTCGGTATGCTG

243/615

**FIGURE 242**

MSGGLPLVLLLTLLGSSHGTGPGMTLQLKLKESFLTNSSESSFLELLEKLCLLLHLPSGTSVTLHHARSQHHVCNT

244/615

**FIGURE 243**

GGCAAGTGGAACCACTGGCTTGGTGGATTTTGCTAGATTTTCTGATTTTTAACTCCTGAAAAATATCCCAGAT  
AACTGTCATGAAGCTGGTAACTATCTTCCTGCTGGTGACCATCAGCCTTTGTAGTTACTCTGCTACTGCCTTCCT  
CATCAACAAAGTGCCCTTCCTGTTGACAAGTTGGCACCTTTACCTCTGGACAACATTCTTCCCTTTATGGATCC  
ATTAAAGCTTCTTCTGAAAACCTCTGGGCATTTCTGTTGAGCACCTTGTGGAGGGGCTAAGGAAGTGTGTAATGA  
GCTGGGACCAGAGGCTTCTGAAGCTGTGAAGAACTGCTGGAGGCGCTATCACACTTGGTGTGACATCAAGATAA  
AGAGCGGAGGTGGATGGGGATGGAAGATGATGCTCCTATCCTCCCTGCCTGAAACCTGTTCTACCAATTATAGAT  
CAAATGCCCTAAAATGTAGTGACCCGTGAAAAGGACAAATAAAGCAATGAATACATTA

245/615

**FIGURE 244**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA59855  
<subunit 1 of 1, 93 aa, 1 stop  
<MW: 10161, pI: 7.39, NX(S/T): 0  
MKLVTFLLVTISLCSYSATAFLINKVPLPVDKLAPLPDNLFPMDPLKLLKTLGISV  
EHLVEGLRKCVNELGPEASEAVKKLLEALSHLV
```

**Important features:****Signal peptide:**

Amino acids 1-18



246/615

**FIGURE 245**

TGCTAGGCTCTGTCCCACAATGCACCCGAGAGCAGGAGCTGAAAGCCTCTAACACCCACAGATCCCTCTATGACT  
GCAATGTGAGGTGTCCGGCTTTGCTGGCCCCAGCAAGCCTGATAAGC**ATGA**AGCTCTTATCTTTGGTGGCTGTGGT  
CGGGTGTTTGCTGGTGCCCCAGCTGAAGCCAACAAGAGTTCTGAAGATATCCGGTGCAAATGCATCTGTCCACC  
TTATAGAAACATCAGTGGGCACATTTACAACCAGAATGTATCCAGAAAGGACTGCAACTGCCTGCACGTGGTGGA  
GCCCATGCCAGTGCCTGGCCATGACGTGGAGGCCTACTGCCTGCTGTGCGAGTGCAGGTACGAGGAGCGCAGCAC  
CACCACCATCAAGGTCATCATTGTCATCTACCTGTCCGTGGTGGGTGCCCTGTTGCTCTACATGGCCTTCCTGAT  
GCTGGTGGACCCTCTGATCCGAAAGCCGGATGCATACACTGAGCAACTGCACAATGAGGAGGAGAATGAGGATGC  
TCGCTCTATGGCAGCAGCTGCTGCATCCCTCGGGGGACCCCGAGCAAACACAGTCCTGGAGCGTGTGGAAGGTGC  
CCAGCAGCGGTGGAAGCTGCAGGTGCAGGAGCAGCGGAAGACAGTCTTCGATCGGCACAAGATGCTCAGC**TAG**AT  
GGGCTGGTGTGGTTGGGTCAAGGCCCCAACACCATGGCTGCCAGCTTCAGGCTGGACAAAGCAGGGGGCTACTT  
CTCCCTTCCTCGGTTCCAGTCTTCCCTTTAAAGCCTGTGGCATTTTTCTCCTTCTCCCTAACTTTAGAAATG  
TTGTACTTGGCTATTTTGATTAGGGAAGAGGGATGTGGTCTCTGATCTCTGTTGTCTTCTTGGGTCTTTGGGGTT  
GAAGGGAGGGGGAAGGCAGGCCAGAAGGGAATGGAGACATTGAGGCGGCCTCAGGAGTGGATGCGATCTGTCTC  
TCCTGGCTCCACTCTTGCCGCCTTCAGCTCTGAGTCTTGGGAATGTTGTTACCCCTTGAAGATAAAGCTGGGTC  
TTCAGGAACCTCAGTGTCTGGGAGGAAAGCATGGCCCAGCATTGAGCATGTGTTCCCTTCTGCAGTGGTTCCTATC  
ACCACCTCCCTCCAGCCCCGCGCCTCAGCCCCAGCCCCAGCTCCAGCCCTGAGGACAGCTCTGATGGGAGAGC  
TGGGCCCCCTGAGCCCACTGGGTCTTCAGGGTGCACTGGAAGCTGGTGTTCGCTGTCCCCTGTGCACTTCTCGCA  
CTGGGGCATGGAGTGCCCATGCATACTCTGCTGCCGGTCCCCTCACCTGCACTTGAGGGGTCTGGGCAGTCCCTC  
CTCTCCCCAGTGTCCACAGTCACTGAGCCAGACGGTTCGGTTGGAACATGAGACTCGAGGCTGAGCGTGGATCTGA  
ACACCACAGCCCCTGTACTTGGGTTCCTCTTGTCCCTGAACCTTCGTTGTACCACTGCATGGAGAGAAAATTTTG  
TCCTCTTGTCTTAGAGTTGTGTGTAAATCAAGGAAGCCATCATTAATTTGTTTTATTTCTCTCA

247/615

**FIGURE 246**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA60278
<subunit 1 of 1, 183 aa, 1 stop
<MW: 20574, pI: 6.60, NX(S/T): 3
MKLLSLVAVVGCLLVPPAEANKSSEDIRCKCICPPYRNISGHIYNQNVSQKDCNCLHVVEPMPVPGHDVEAYCLI
CECRYEERSTTTIKVIIIVYLSVVGALLLYMAFLMLVDPLIRKPDAYTEQLHNEEENEDARSMAAAAASLGGPRA
NTVLERVEGAQQRWKLQVQEQRKTVFDRHKMLS
```

**Important features:****Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 90-112

**N-glycosylation sites.**

amino acids 21-24, 38-41 and 47-50

248/615

**FIGURE 247**

AATTGTATCTGTGTAATGTTAAAAACAAACGAAATAAAATAGAAGGAAAAACTTTCTGAGTTTCAAAAACAACAGA  
CTAGTACTCTAAAGAACTCTTTAAAAACAATTAAGTGTAGGATTGCAGTTATGATTGGATATTATTTAATTCTGT  
TTCTGATGTGGGGTTCCCTCCACTGTGTTCTGTGTGCTATTAATATTTACCATTGCAGAAGCTTCATTCAGTGTG  
AAAATGAATGCTTAGTGGATCTGTGCCTCTTACGCATATGTTACAAATTATCTGGAGTTCCTAATCAATGCAGAG  
TTCCCTCCCTCCGATTGTTCTAAATTAATTGAAAGATGTCTGCTGTGGAAAAAGGCATGTATTTAAATCTGTAT  
GATTCTCAACCATCTTTAGTTGGGAAAGGTCCTTGAAAGCCAATGGAAATACTTTTTTTTTTTCTTGGCACTAAT  
CAAGTGAGTGTTACCTTTTCACTTAGTAGGATGTGTTGTTACGCTAGTAAAATAGAAACCTGTGTTTATTCTCAG  
GTATTTTAGAAACAACAGCCATCATTTTATTTTATGTGTGTGTTCTTGGCTGTATTCATAAATTATATATTTTGG  
GCTATCAAATATTACTTCATTCAATATAAATAACAATAGTAGAAGTTGTTTACTTAGATATGCTTTCTAGTTGCA  
TTTTCTCAGCCTATGTAAGACTACTTTGTTGTAATAGCCTTTGAAATTACAGTACTGTCTCTCTACTATCTTCA  
GATTACTTGATTCAAATAAACCAATTATGTTTGTAATTGATATTAATAAAACCAGAATAAAAGTTTCATATCTACCC

249/615

**FIGURE 248**

MIGYYLILFLMWGSSTVFCVLLIFTIAEASFVENECLVDLCLLRICYKLSGVPNQCRVPLPSDCSK

**Important features:**

**Signal peptide:**

amino acids 1-29

250/615

**FIGURE 249**

AGCGGGTCTCGCTTGGGTTCGGCTAATTTCTGTCCTGAGGCGTGAGACTGAGTTCATAGGGTCCTGGGTCCCCGA  
ACCAGGAAGGGTTGAGGGAACACAATCTGCAAGCCCCCGCAGCCCAAGTGAGGGGGCCCCGTGTTGGGGTCCTCCC  
TCCCTTTGCATTCCCACCCCTCCGGGCTTTGCGTCTTCTGGGGACCCCTCGCCGGGAGATGCGCCGCTTGATG  
CGGAGCAAGGATTCGTCTGCTGCTGCTCCTACTGGCCGCGGTGCTGATGGTGGAGAGCTCACAGATCGGCAGT  
TCGCGGGGCCAACTCAACTCCATCAAGTCTCTCTGGGCGGGGAGACGCTGGTCAGGCCGCCAATCGATCTGCG  
GGCATGTACCAAGGACTGGCATTGCGCGGCAGTAAGAAGGGCAAAAACCTGGGGCAGGCCTACCCTTGTAGCAGT  
GATAAGGAGTGTGAAGTTGGGAGGTATTGCCACAGTCCCCACCAAGGATCATCGGCCTGCATGGTGTGTCGGAGA  
AAAAAGAAGCGCTGCCACCGAGATGGCATGTGCTGCCCCAGTACCCGCTGCAATAATGGCATCTGTATCCCAGTT  
ACTGAAAGCATCTTAACCCCTCACATCCCGGCTCTGGATGGTACTCGGCACAGAGATCGAAACCACGGTCATTAC  
TCAAACCATGACTTGGGATGGCAGAATCTAGGAAGACCACACACTAAGATGTCACATATAAAAGGGCATGAAGGA  
GACCCCTGCCTACGATCATCAGACTGCATTGAAGGGTTTTGCTGTGCTCGTCATTTCTGGACCAAAATCTGCAAA  
CCAGTGCTCCATCAGGGGGAAGTCTGTACCAAAACAACGCAAGAAGGGTTCTCATGGGCTGGAAATTTTCCAGCGT  
TGCGACTGTGCGAAGGGCCTGTCTTGCAAAGTATGGAAAGATGCCACCTACTCCTCAAAGCCAGACTCCATGTG  
TGTCAGAAAATTTGATCACCATTGAGGAACATCATCAATTGCAGACTGTGAAGTTGTGATTTAATGCATTATAG  
CATGGTGGAAAATAAGGTTCCAGATGCAGAAGAATGGCTAAAATAAGAAACGTGATAAGAATATAGATGATCACA  
AAAGGGAGAAAGAAAACATGAAGTGAATAGATTAGAATGGGTGACAAATGCAGTGCAGCCAGTGTTCATTATG  
CAACTTGTCTATGTAAATAATGTACACATTTGTGGAAAATGCTATTATTAAGAGAACAAGCACACAGTGGAAAT  
ACTGATGAGTAGCATGTGACTTTCCAAGAGTTTAGGTTGTGCTGGAGGAGAGGTTTCCTTCAGATTGCTGATTGC  
TTATACAAATAACCTACATGCCAGATTTCTATTCAACGTTAGAGTTTAAACAAAATACTCCTAGAATAACTTGTTA  
TACAATAGGTTCTAAAAATAAAATTGCTAAACAAGAAATGAAAACATGGAGCATTTGTTAATTTACAACAGAAAAT  
TACCTTTTGATTTGTAACACTACTTCTGCTGTTCAATCAAGAGTCTTGCTAGATAAGAAAAAATCAGTCAATAT  
TTCCAAATAATTGCAAAATAATGGCCAGTTGTTTAGGAAGGCCCTTAGGAAGACAAATAAATAACAACAAACAG  
CCACAAATACTTTTTTTTCAAATTTTAGTTTTACCTGTAATTAATAAGAACTGATACAAGACAAAACAGTTCC  
TTCAGATTCTACGGAATGACAGTATATCTCTTTATCCTATGTGATTCTGCTGCTGAATGCATTATATTTTCCA  
AACTATACCCATAAATTGTGACTAGTAAAATACTTACACAGAGCAGAATTTTTCAGATGGCAAAAAAATTTAAA  
GATGTCCAATATATGTGGGAAAAGAGCTAACAGAGAGATCATTATTTCTTAAAGATTGGCCATAACCTATATTTT  
GATAGAATTAGATTGGTAAATACATGTATTACATACATACTCTGTGGTAATAGAGACTTAAGCTGGATCTGTACTG  
CACTGGAGTAAGCAAGAAAATTGGGAAAACTTTTTCGTTTGTTCAGGTTTTGGCAACACATAGATCATATGCTG  
AGGCACAAGTTGGCTGTTTCATCTTTGAAACCAGGGGATGCACAGTCTAAATGAATATCTGCATGGGATTTGCTAT  
CATAATATTTACTATGCAGATGAATTCAGTGTGAGGTCTGTGTCCGTACTATCCTCAAATTATTTATTTTATAG  
TGCTGAGATCCTCAAATAATCTCAATTTCAGGAGGTTTTACAAAATGTACTCCTGAAGTAGACAGAGTAGTGAGG  
TTTCATTGCCCTCTATAAGCTTCTGACTAGCCAATGGCATCATCCAATTTTCTTCCCAAACCTCTGCAGCATCTG  
CTTTATTGCCAAAGGGCTAGTTTCGGTTTTCTGCAGCCATTGCGGTTAAAAAATATAAGTAGGATAACTTGTAAA  
ACCTGCATATTGCTAATCTATAGACACCACAGTTTCTAAATTTCTTGAACCACTTTACTACTTTTTTTAACTT  
AACTCAGTTCTAAATACTTTGTCTGGAGCACAAAACAATAAAAGGTTATCTTATAGTCGTGACTTTAACTTTTG  
TAGACCACAATTCACTTTTTAGTTTTCTTTTACTTAAATCCCATCTGCAGTCTCAAATTTAAGTTCTCCAGTAG  
AGATTGAGTTTGAGCCTGTATATCTATTAAAAATTTCAACTTCCACATATATTTACTAAGATGATTAAGACTTA  
CATTTTCTGCACAGGTCTGCAAAAACAAAATTATAAACTAGTCCATCCAAGAACCAAGTTTGTATAAACAGGT  
TGCTATAAGCTTGTGAAATGAAAATGGAACATTTCAATCAAACATTTCTATATAACAATTATTATTTACAAT  
TTGGTTTCTGCAATATTTTTCTTATGTCCACCCTTTAAAAATTATTATTTGAAGTAATTTATTTACAGGAAATG  
TTTGTAAATTTAGGAAAAATGTATAATAAGATAAAATCTATTAAATTTTTCTCCTCTAAAAACTGAAAAA  
AAAAA

251/615

**FIGURE 250**

MAALMRSKDSSCCLLLLAAVLMVESSQIGSSRAKLNSIKSSLGGETPGQAANRSAGMYQGLAFGGSKKGKNLGQA  
YPCSSDKECEVGRYCHSPHQSSACMVCRRKKKRCHRDGMCCPSTRCNGICIPVTESILTPHIPALDGTTRHRDR  
NHGHYSNHDLGWQNLGRPHTKMSHIKGHEGDPCLRSSDCIEGFCCARHEFTKICKPVLHQGEVCTKQRKKGSHGL  
EIFQRCDCAKGLSCKVWKDATYSSKARLHVCQKI

**Signal peptide:**  
amino acids 1-25

252/615

**FIGURE 251**

TCTCAATCTGCTGACCTCGTGATCCGCGCTGACCTTGTAATCCACCTACCTTGGCCTCCCAAAGTGTTGGGATTAC  
AGGCGTGAGCCACCGCGCCCGGCCAACATCACGTTTTTAAAAATTGATTTCTTCAAATTCATGGCAAATATTTCC  
CTTCCCTTTAACTTCTTATGTCAGAATGAGGAAGGATAGCTGCATTTATTTAGTCAGTTTTTCATTGCATAGTAAT  
ATTTTCATGTAGTATTTTCTAAGTTATATTTTAGTAATTCATATGTTTTTAGATTATAGGTTTTAACATACTTGTG  
AAAATACTTGATGTTTTAAAGCCTTGGGCAGAAATTCGTATTGTTGAGGATTTGTTCTTTTATCCCCCTTTT  
AAAGTCATCCGTCCTTGGCTCAGGATTTGGAGAGCTTGCAACCAAAAAATGGCAAACATCACCAGCTCCAGAT  
TTTGGACCAGTTGAAAGCTCCGAGTTTGGGCCAGTTTACCACCACCCCAAGTACACAGCAGAATAGTACAAGTCA  
CCCTACAATACTACTTCTTGGGACCTCAAGCCCCCAACATCCCAGTCCTCAGTCCTCAGTCATCTTGACTTCAA  
ATCTCAACCTGAGCCATCCCCAGTTCTTAGCCAGTTGAGCCAGCGACAACAGCACCAGAGCCAGGCAGTCACGTGT  
TCCTCCTCCTGGTTTGGAGTCCTTTCCCTTCCCAGGCAAACTTCGAGAATCAACACCTGGAGACAGTCCCTCCAC  
TGTGAACAAGCTTTTGCAGCTTCCCAGCACGACCATTTGAAAAATATCTCTGTGTCTGTCCACCAGCCACAGCCCAA  
ACACATCAAACCTTGCTAAGCGGCGGATACCCCCAGCTTCTAAGATCCCAGCTTCTGCAGTGGAAATGCCTGGTTC  
AGCAGATGTCACAGGATTAAATGTGCAGTTTGGGGCTCTGGAATTTGGGTGAGAACCTTCTCTCTCTGAATTTGG  
ATCAGCTCCAAGCAGTGAAAATAGTAATCAGATTCCCATCAGCTTGTATTGGAAGTCTTTAAGTGAGCCTTTGAA  
TACATCTTTATCAATGACCAGTGCAGTACAGAACTCCACATATACAACTTCCGTCATTACCTCCTGCAGTCTGAC  
AAGCTCATCACTGAATTCTGCTAGTCCAGTAGCAATGTCTTCTCTTATGACCAGAGTTCTGTGCATAACAGGAT  
CCCATACCAAAGCCCTGTGAGTTCATCAGAGTCAGCTCCAGGAACCATCATGAATGGACATGGTGGTGGTCTGAA  
TCAGCAGACACTAGACAGTAAGTATAGCAGCAAGCTACTCTTGTCTATGGCTGGTGCCAACCAAACAGAGGAAGAG  
GATAGCTCACGTGATGTGGAAAACACCAGTTGGTCAATGGCTCATTCGTTAAAAGCAGCCCTTTTGCTTTTTTG  
TTTTTGGACCAGGTGTTGGCTGTGGTGTATTAGAAATGTCTTAACCACAGCAAGAAGGAGGTGGTGGTCTCATA  
TTCTTCTGCCCTAATCAGACTGCACCACAAGTGCAGCATACAGTATGCATTTTAAAGATGCTTGGGCCAGGCGGG  
GTGGCTGATGCCATAATCCCAGTGCTTTGGGGGGCCAAGGCAGGCAGATTGCCCAAGCTCAGGAGTTTGAGACC  
ACCCTGGGCAACATGGTGAAACTCTGTCTCTACTAAAAATACGAAAACTAGCCGGGTGTGGTGGCGCGCGCTGCC  
TGTAATCCCAGCTACTTGGGAGGCTGAGGCACAAGAATCGCTTGAGCCAGCTTGGGCTACAAAGTGAGACTCCGT  
CTGAAAGA

253/615

**FIGURE 252**

MCFKALGRNSVLLRICSFIPLLKSSVLGSGFGELAPPKMANITSSQILDQLKAPSLGQFTTTPSTQQNSTSHPTT  
TTSWDLKPPTSQSSVLSHLDFKSQPEPSPVLSQLSQRQQHQSQAQVTVPPPGLESFPSQAKLRESTPGDSPSTVVK  
LLQLPSTTIENISVSVHQPQPKHIKLAQRRIPPASKIPASAVEMPGSADVTGLNVQFGALEFGSEPSLSEFGSAP  
SSENSNQIPISLYSKSLSEPLNTSLSMSTSAVQNSTYTTSVITSCSLTSSSLNSASPVAMSSSYDQSSVHNRIPIYQ  
SPVSSSESAPGTIMNGHGGGRSQQTLDISKYSSKLLLSWLVPVKQRKRIAHVMWKTPVGQWLIR

**Signal peptide:**  
amino acids 1-24



**FIGURE 253**

[illegible]

255/615

**FIGURE 254**

MELSQMSELMGLSVLLGLLALMATAAVARGWLRAGEERSGRPACQKANGFPPDKSSGSKKQKQYQIRIRKEKPQQH  
NFTHRLAAALKSHSGNISCMDFSSNGKYLATCADDRTIRIWSTKDFLQREHRSMRANVELDHATLVRFSPDCRA  
FIVWLANGDTLRVFKMTKREDGGYTFTATPEDFPKKHKAPVIDIGIANTGKFIMTASSDTTVLIWSLKGQVLSTI  
NTNQMNNTHAAVSPCGRFVASCFTPDVKVWEVCFGKKGEFQEVVRAFELKGHSAAVHSFAFSNDSRRMASVSKD  
GTWKLWDTDVEYKKKQDPYLLKTGRFEEAAGAAPCRLALSPNAQVLALASGSSIHLYNTRRGEKEECFERVHGEC  
IANLSFDITGRFLASCGDRAVRLFHNTPGHRAMVEEMQGHKLRASNESTRQRLQQQLTQAQETLKSLGALKK

**Important features:****Signal peptide:**

amino acids 1-25

**N-glycosylation site.**

amino acids 76-80, 92-96, 231-235, 289-293, 378-382, 421-425

**Beta-transducin family Trp-Asp repeat protein.**

amino acids 30-47, 105-118, 107-119, 203-216, 205-217, 296-308

256/615

**FIGURE 255**

ACGGACCGAGGGTTCGAGGGAGGGACACGGACCAGGAACCTGAGCTAGGTCAAAGACGCCCCGGGCCAGGTGCCCC  
GTCGCAGGTGCCCCCTGGCCGGAGATGCGGTAGGAGGGGCGAGCGGAGAAGCCCCCTTCCTCGGCGCTGCCAACCC  
GCCACCCAGCCC**ATG**GCGAACCCCGGGCTGGGGCTGCTTCTGGCGCTGGGCCTGCCGTTCTGCTGGCCCGCTGG  
GGCCGAGCCTGGGGGCAAATACAGACCACTTCTGCAAATGAGAATAGCACTGTTTGCCTTCATCCACCAGCTCC  
AGCTCCGATGGCAACCTGCGTCCGGAAGCCATCACTGCTATCATCGTGGTCTTCTCCCTCTTGGCTGCCTTGCTC  
CTGGCTGTGGGGCTGGCACTGTTGGTGCGGAAGCTTCGGGAGAAGCGGCAGACGGAGGGCACCTACCGGCCAGT  
AGCGAGGAGCAGTTCTCCCATGCAGCCGAGGCCCGGGCCCCCTCAGGACTCCAAGGAGACGGTGCAGGGCTGCCTG  
CCCATC**TAG**GTCCCCTCTCCTGCATCTGTCTCCCTTCATTGCTGTGTGACCTTGGGGAAAGGCAGTGCCCTCTCT  
GGGCAGTCAGATCCACCCAGTGCTTAATAGCAGGGAAGAAGGTACTTCAAAGACTCTGCCCCTGAGGTCAAGAGA  
GGATGGGGCTATTCACTTTTATATATTTATATAAAATTAGTAGTGAGATGTAAAAAAAAAAAAAAAAAAAA

257/615

**FIGURE 256**

MANPGLGLLLALGLPFLARWGRAWGQIQTTSANENSTVLPSTSSSSDGNLRPEAITAIIVVFSLLAALLAVG  
LALLVRKLRKREQTEGTYRPSSEEQFSHAAEARAPQDSKETVQGCLPI

**FIGURE 257**

GCCAGGAATAACTAGAGAGGAACAATCGGGTTATTTCAGAGGTTTTGTTTTCTCTTAGTCTGTGCGCTGCTGCAC  
CAGTCAAAATCTTCTCTTCAATTAAGCTGAATAATAATGGCTTTGAAGATATTGTCAATTGTTATAGATCCTAGTGTG  
CCAGAAAGATGAAAAATAATTAACAAATAGAGGATATGTTGACTACAGCTCTACGTTACCTGTTGAAGCCACA  
GAAAAAAGATTTTTTTTTCAAAAATGTATCTATATTAATCTCGTAAGATTGGAAGGAAAACTCAGTACAAAAGG  
CCAAAACATGAAACCATAAACATGCTGATGTTATAGTTGCAACCCTACACTCCCAGGTAGAGATGAACCATAC  
ACCAAGCAGTTTCAGAAATGTGGAGAGAAAGCGAATACATTCATCTACCCTGACCTTCTACTTGGAAAAAAA  
CAAAATGAATATGGACCACCAGGCAAACCTGTTGTCCATGAGTGGGCTCACCTCCGGTGGGGAGTGTGTTGATGAG  
TACAATGAAGATCAGCCTTTCTACCGTGCTAAGTCAAAAAAATCGAAGCAACAAGGTGTTCCCGCAGGTATCTCT  
GGTAGAAATAGAGTTTTATAAGTGTCAAGGAGGCAGCTGTCTTAGTAGAGCATGCAGAAATGATTCTACAACAAAA  
CTGTATGGAAGAAGATTGTCAATCTTTCTGATAAAGTACAAACGAAAAAGCATCCATAATGTTTATGCAAAAGT  
ATTGATTTCTGTTTGAATTTTGTAAAGAAAAAACCCATAATCAGAAGCTCCAGGCTACAAAACATAAAGTGC  
AATTTTGAAGATCATGGGAGGTGATTGCAAAATCTGAGGATTTTAAAAACCACTACCCATGGTGACACCCT  
CCTCCACCTGCTTCTCATTGCTGAAGATCAGTCAAAGAATTGTGTGCTTAGTCTTGATAAGTCTGGAAGCATG  
GGGGTGAAGGACCGCCTAATCGAATGAATCAAGCAGCAAAACATTTCTGCTGACAGCTGTTGAAATGGATGC  
TGGTGGGGATGTTTCACTTTGATAGTACTGCCACTATTGTAAATAAGCTAATCAAATAAAAGACAGTGTGAA  
AGAAACACTCATGGCAGGATTACCTACATATCCTCTGGGAGGAACTCCATCTGCTCTGGAATTAATATGCA  
TTTCAGGTGATTGGAGAGCTACATTTCCCACTCGATGGATCCGAAGTCTGCTGCTGACTGATGGGAGGATAAC  
ACTGCAAGTTCTTTGATTGATGAGTGAAGTGAACAAAGTGGGGCCATTGTTCATTTGATTGTTTTGGGAAGAGCTGCT  
GATGAAGCAGTAAATAGAGTAGAGTAGCAAGATAACAGGAGGAAGTCATTTTTATGTTTCAGATGAAGCTCAGAACAAT  
GGCTCATTGATGCTTTTTGGGGCTCTTACATCAGGAATACTGATCTCTCCGAAGTCCCTTCAGCTCGAAAGT  
AAGGGATTAACACTGAATAGTAATGCCTGGATGAACGACACTGTCATAATTGATAGTACAGTGGGAAGGACAG  
TTCTTTCTCATCATCATGGAACAGTCTGCCTCCCAGTATTTCTCTCTGGGATCCCAGTGAACAATAATGGAAAA  
TTCACAGTGGATGCAACTTCCAAATGGCCTATCTCAGTATTCCAGGAACGTGCAAGGCTGGGCATTTGGGCATAC  
AATCTTCAAGCCAAAGCGAAGCAACCAAGAACTAATCTACAGTAACCTCTCGAGCAGAAATCTTCTCTGCTCT  
CCAATCAGAGTGAATGCTAAAATGAATAAGGACGTAAACAGTTTCCCCAGCCCAATGATTGTTTACGCAGAAAT  
CTACAAGGATATGTACCTTCTTGGAGCCAATGTGACTGCTTTCATTGAATCAGAAATGGACATCAGAAAGTT  
TTGAACTTTTGGATAATGTTGCGAGGCGCTGATTTCTTCAAGATGATGGAGTCTACTCCAGGTATTTTACAGCA  
TATACAGAAAATGGCAGATATAGCTTAAAGTTTGGGCTCATGGAGGAGCAACACTGCCAGGCTAAAATTACGG  
CCTCCAATGAATAGAGCGCGGTACATACAGGCTGGGTAGTGAACGGGGAATTTGAAGCAAAACCGCCCAAGCTT  
GAAATTGATGAGGATACTCAGACACCTTGGAGGATTTCAAGCAACAGCATCCGGAGGTGCATTTGTGTTATCA  
CAAGTCCCCAGCTTCCCTTGCTTGCTGACCAATCCCACCAAGTCAAAATCAGACAGTTGATGCCACAGTTTCATGAG  
GATAAGATTATTTCTACATGGCAGCAGCAGGAGATAATTTGATGTTGGAAAAAGTTCAACGTTTATCATAGA  
ATAAGTGCAAGTATTTCTGATCTAAGAGACAGTTTTGATGATGCTCTTCAAGTAAATACTACTGATCTGTCA  
AAGGAGGCCAACTCCAAGGAAGCTTGCATTTTAAACAGAAAATATCTCAGAAGAAAATGCAACCCACATATTT  
ATTGCCATTTAAAGTATAGATAAAGCAATTTGACATCAAAGTATCCAACTTGACAAAGTAACTTTGTTTATC  
CCTCAAGCAAACTCTGATGACATTGATCCTACACCTACTCCTACTCCTACTCCTACTCCTGATAAAGTCATAAT  
TCTGGAGTTAATATTTCTACGCTGGTATTGCTGTGATTTGGGCTGTGTTGTAATTTGTTAACTTTTAAAGTACC  
ACCATTTGAACCTTAAACGAAGAAAAAATCTCAAGTAGACCTAGAAGGAGTTTAAAAAACAAAAAGATGTA  
GTAAAGGATATTTCTGAATCTTAAATTCATCCATGTGTGATCATAAACTCATAAAAAATAATTTTAAAGTGTG  
GAAAAGGATACCTTGTATTAATAAAACCTCATGGATATGTAAAAAGTGTCAAGATTAATAATTTAATAGTTTCA  
TTTATTTGTTATTTTGTGTAAGAAATAGTGATGAACAAAGATCCTTTTTCATACTGATACCTGGTTGTATAT  
ATTTGATGCAACAGTTTCTGAAATGATATTCAAATTCATCAAGAAATTAATAATCATCTATCTGAGTAGTCAA  
AATACAAGTAAGGAGAGCAATAAACACATTTGGAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA  
AAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA

259/615

**FIGURE 258**

MGLFRGFVFLVLCLLHQSNTSFIKLNNGFEDIVIVIDPSVPEDEKIIIEQIEDMVTASTYLFEATEKRFFFKN  
VSILIPENWKENPQYKRPKHENHKHADVIVAPPTLPGRDEPYTKQFTECGEKGEYIHFTPDLLLGKKQNEYGPPG  
KLFVHEWAHLRWGVFDEYNEDQPFYRAKSKKIEATRCASAGISGRNRVYKCQGGSCLSRACRIDSTTKLYGKDCQF  
FPDKVQTEKASIMFMQSIDSVVEFCNEKTHNQEAPSLQNIKCNRSTWEVISNSED FKNTIPMVT P P P P P V F S L L  
KISQRIVCLVLDDKSGSMGGKDR LNR MNQA AKH FLLQT VENG SWVGMVHFDSTATIVNKLIQIKSSDERNTLMAGL  
PTYPLGGTSICSGIKYAFQVIGELHSQLDGSEVLLLT DGEDNTASSCIDEVKQSGAIVHFIALGRAADEAVIEMS  
KITGGS HFYVSDEAQNNGLIDAFGALTSGNTDLSQKSLQLESKGLT LNSNAWMNDTVIIDSTVGKDTFFLITWNS  
LPPSISLWDPSGTIMENFTVDATSKMAYLSIPGTAKVGTWAYNLQAKANPETLTITVT SRAANSSVPPITVNAKM  
NKDVNSFPSPMIVYAEILQGYVPVLGANVTAFIESQNGHTEVLELLDNGAGADSFKN DGVYSRYFTAYTENGRYS  
LKVRAHGGANTARLKL RPPLNRAAYIPGWV VNGEIEANPPRPEIDEDTQT TLEDFSR TASGGA FVVSQVPSLPLP  
DQYPPSQITDL DATVHEDKIILTWTAPGDNFDVGKVQRYIIRISASILDLRDSFDDALQVNTTDLSPKEANSKES  
FAFKPENISEENATHIFIAIKSIDKSNLTSKVS NIAQVTLFIPQANPDDIDPTPTPTPTPTPKSHNSGVNISTL  
VLSVIGSVVIVNFILSTTI

**Signal peptide:**

amino acids 1-21

**Putative transmembrane domains:**

amino acids 284-300, 617-633

**Leucine zipper pattern.**

amino acids 469-491, 476-498

**N-glycosylation site.**amino acids 20-24, 75-79, 340-344, 504-508, 542-546, 588-592, 628-632,  
811-815, 832-836, 837-841, 852-856, 896-900

260/615

**FIGURE 259**

CGCCGGAGGCAGCGGGCGGCGTGGCGCAGCGGCGACATGGCCGTTGTCTCAGAGGACGACTTTCAGCACAGTTCAA  
ACTCCACCTACGGAACCACAAGCAGCAGTCTCCGAGCTGACCAGGAGGCACTGCTTGAGAAGCTGCTGGACCGCC  
CGCCCCCTGGCCTGCAGAGGCCCCGAGGACCGCTTCTGTGGCACATACATCATCTTCTTCAGCCTGGGCATTGGCA  
GTCTACTGCCATGGAACCTCTTTATCACTGCCAAGGAGTACTGGATGTTCAAACCTCCGCAACTCCTCCAGCCCAG  
CCACCGGGGAGGACCCCTGAGGGCTCAGACATCCTGAACACTTTTGAGAGCTACCTTGCCGTTGCCTCCACCGTGC  
CCTCCATGCTGTGCCTGGTGCCAACTTCTGCTTGTCAACAGGGTTGCAGTCCACATCCGTGTCTGGCCCTCAC  
TGACGGTCATCCTGGCCATCTTCATGGTGATAACTGCACTGGTGAAGGTGGACACTTCTCCTGGACCCGTGGTT  
TTTTTGCGGTACCAATTGTCTGCATGGTGATCCTCAGCGGTGCCTCCACTGTCTTCAGCAGCAGCATCTACGGCA  
TGACCGGCTCCTTTCTATGAGGAACTCCCAAGCACTGATATCAGGAGGAGCCATGGGCGGGACGGTCAGCGCCG  
TGGCCTCATTGGTGGACTTGGCTGCATCCAGTGATGTGAGGAACAGCGCCCTGGCCTTCTTCTGACGGCCACCA  
TCTTCCTCGTGCTCTGCATGGGACTCTACCTGCTGCTGTCCAGGCTGGAGTATGCCAGGTACTACATGAGGCCCTG  
TTCTTGCGGCCCCATGTGTTTTCTGGTGAAGAGGAGCTTCCCCAGGACTCCCTCAGTGCCCTTCGGTGGCCCTCCA  
GATTCAATTGATTCCCACACACCCCCCTCTCCGCCCCATCCTGAAGAAGACGGCCAGCCTGGGCTTCTGTGTACCT  
ACGTCTTCTTCATCACCAGCCTCATCTACCCCGCCGTCTGCACCAACATCGAGTCCCTCAACAAGGGCTCGGGCT  
CACTGTGGACCACCAAGTTTTTCATCCCCCTCACTACCTTCTCCTGTACAACTTTGCTGACCTATGTGGCCGGC  
AGCTCACCGCCTGGATCCAGGTGCCAGGGCCCCAACAGCAAGGCGCTCCAGGGTTCGTGCTCCTCCGACCTGCC  
TCATCCCCCTCTTCGTGCTCTGTAACCTACCAGCCCCGCGTCCACCTGAAGACTGTGGTCTTCCAGTCCGATGTGT  
ACCCGCACTCCTCAGCTCCCTGCTGGGGCTCAGCAACGGCTACCTCAGCACCCCTGGCCCTCCTCTACGGGCCCTA  
AGATTGTGCCCAGGGAGCTGGCTGAGGCCACGGGAGTGGTGATGTCTTTTATGTGTGCTTGGGCTTAACACTGG  
GCTCAGCCTGCTCTACCTCCTGGTGACCTCATCTAGAAAGGGAGGACACAAGGACATTGGTGCTTCAGAGCCTT  
TGAAGATGAGAAGAGAGTGCAGGAGGGCTGGGGGCCATGGAGGAAAGGCCTAAAGTTTCACTTGGGGACAGAGAG  
CAGAGCACACTCGGGCCTCATCCCTCCCAAGATGCCAGTGAGCCACGTCCATGCCCATTCCGTGCAAGGCAGATA  
TTCCAGTCATATTAAACAGAACACTCCTGAGACAGTTGAAGAAGAAATAGCACAAATCAGGGGTACTCCCTTCACA  
GCTGATGGTTAACATTCCACCTTCTTTCTAGCCCTTCAAAGATGCTGCCAGTGTTCGCCCTAGAGTTATTACAAA  
GCCAGTGCCAAAACCCAGCCATGGGCTCTTTGCAACCTCCCAGCTGCGCTCATTCCAGCTGACAGCGAGATGCAA  
GCAAATGCTCAGCTCTCCTTACCCTGAAGGGGTCTCCCTGGAATGGAAGTCCCCTGGCATGGTCAGTCTCAGGC  
CCAAGACTCAAGTGTGCACAGACCCCTGTGTTCTGCGGGTGAACAACCTGCCCCTAACCAGACTGGAAAACCCAG  
AAAGATGGGCTTCCATGAATGCTTCATTCCAGAGGGACCAGAGGGCCTCCCTGTGCAAGGGATCAAGCATGTCT  
GGCCTGGGTTTTTCAAAAAAAGAGGGATCCTCATGACCTGGTGGTCTATGGCCTGGGTCAAGATGAGGGTCTTTCA  
TGTTCTCTGTTTACAACATGTCAAAGCCATTGGTTCAAGGGCGTAATAAATACTTGCGTATTCAAAA

261/615

**FIGURE 260**

MAVVS EDDFQHSSNSTYGTSSSLRADQEALLEKLLDRPPPGQLQRPEDRFCGTYYIIFSLGIGSLLPWNFFITAK  
EYWMFKLRNSSSPATGEDPEGSDILNYFESYLAVASTVPSMLCLVANFLLVNRVAVHIRVLASLTVILAI FMVIT  
ALVKVDTSSWTRGFFAVTIVCMVILSGASTVFSSSIYGMTGSFPMRNSQALISGGAMGGTVSAVASLVDLAASSD  
VRNSALAFFLTATIFLVLCMGLYLLLSRLEYARYYMRPVLAHVFSGEEELPQDSL SAPSVASRFIDSHTPPLRP  
ILKKTASLGFCVTYVFFITS LIYPAVCTNIESLNKSGSGLWTTKFFIPLTTFLLYNFADLCGRQLTAWIQVPGPN  
SKALPGFVLLRTCLIPLFVLCNYQPRVHLKTVVFQSDVYPALLSSLLGLSNGYLSTLALLYGPKIVPRELAEATG  
VVMSFYVCLGLTLGSACSTLLVHLI

**Transmembrane domain:**

amino acids 50-74 (type II), 105-127, 135-153, 163-183, 228-252, 305-330,  
448-472



262/615

**FIGURE 261**

CGGACGCGTGGGCTGCTGGTGGGAAGGCCTAAAGAACTGGAAAGCCCCTCTCTTGGAAACCACACCTGTTTTA  
AAGAACCTAAGCACCATTAAAGCCACTGGAAATTTGTTGTCTAGTGGTTGTGGGTGAATAAAGGAGGGCAGAAAT  
GGATGATTTTCATCTCCATTAGCCTGCTGTCTCTGGCTATGTTGGTGGGATGTTACGTGGCCGGAATCATTCCCTT  
GGCTGTTAATTTCTCAGAGGAACGACTGAAGCTGGTGACTGTTTTGGGTGCTGGCCTTCTCTGTGGAACCTGCTCT  
GGCAGTCATCGTGCCCTGAAGGAGTACATGCCCTTTATGAAGATATTCTTGAGGGAAAAACACCACCAAGCAAGTGA  
AACACATAATGTGATTGCATCAGACAAAGCAGCAGAAAAATCAGTTGTCCATGAACATGAGCACAGCCACGACCA  
CACACAGCTGCATGCCCTATATTGGTGTTCCTCGTTCTGGGCTTCGTTTTTCATGTTGCTGGTGGACCAGATTGG  
TAACTCCCATGTGCATTCTACTGACGATCCAGAAGCAGCAAGGTCTAGCAATCCAAAATCACCACCACGCTGGG  
TCTGGTTGTCCATGCTGCAGCTGATGGTGTTCCTTGGGAGCAGCAGCATCTACTTCACAGACCAGTGTCCAGTT  
AATTGTGTTTGTGGCAATCATGCTACATAAGGCACCAGCTGCTTTTGGACTGGTTTCCTTCTTGATGCATGCTGG  
CTTAGAGCGGAATCGAATCAGAAAGCACTTGCTGGTCTTTGCATTGGCAGCACCAGTTATGTCCATGGTGACATA  
CTTAGGACTGAGTAAGAGCAGTAAAGAAGCCCTTTCAGAGGTGAACGCCACGGGAGTGGCCATGCTTTTCTCTGC  
CGGGACATTTCTTTATGTTGCCACAGTACATGTCTCCCTGAGGTGGGCGGAATAGGGCAGCCACAAGCCCGA  
TGCCACGGGAGGGAGAGGCTCAGCCGCTGGAAGTGGCAGCCCTGGTTCTGGGTTGCCTCATCCCTCTCATCCT  
GTCAGTAGGACACCAGCATTAATGTTCAAGGTCCAGCCTTGGTCCAGGGCCGTTGCCATCCAGTGAGAACAGC  
CGGCACGTGACAGCTACTCACTTCCTCAGTCTCTTGTCTCACCTTGCGCATCTCTACATGTATTCTAGAGTCCA  
GAGGGGAGGTGAGGTTAAACCTGAGTAATGGAAAAGCTTTTAGAGTAGAAACACATTTACGTTGCAGTTAGCTA  
TAGACATCCCATTTGTGTTATCTTTTAAAGGCCCTTGACATTTTGCCTTTTAAATATTTCTCTTAACCTATTCTC  
AGGGAAGATGGAAATTTAGTTTTAAGGAAAAGAGGAGAACTTCATACTACAATGAAATAGTGATTATGAAAATAC  
AGTGTCTGTAAATTAAGCTATGTCTCTTTCTTCTAGTTTAGAGGCTCTGCTACTTTATCCATTGATTTTTTAACA  
TGGTTCACCACATGTAAGACTGGTGCTTTAGCATCTATGCCACATGCGTTGATGGAAGGTGATAGCACCACCTCA  
CTTAGATGCTAAAGGTGATTCTAGTTAATCTGGGATTAGGGTCAGGAAAATGATAGCAAGACACATTGAAAGCTC  
TCTTTATACTCAAAAGAGATATCCATTGAAAAGGGATGTCTAGAGGGATTTAAACAGCTCCTTTGGCAGCTGCCCT  
CTCTGAATCCAGCCTGCCATTCCATCAAATGGAGCAGGAGAGGTGGGAGGAGCTTCTAAAGAGGTGACTGGTATT  
TTGTAGCATTCCCTTGTCAAGTTCTCCTTTGCAGAATACCTGTCTCCACATTCCTAGAGAGGAGCCAAGTCTAGT  
AGTTTCAGTTCTAGGCTTTCCTTCAAGAACAGTCAGATCACAAAGTGTCTTTGGAAATTAAGGGATATTAAATTT  
TAAGTGATTTTTGGATGGTTATTGATATCTTTGTAGTAGCTTTTTTTTAAAGACTACCAAAATGTATGGTTGTCC  
TTTTTTTTTGTTTTTTTTTTTTTTAAATATTTCTCTTAGCAGATCAGCAATCCCTCTAGGGACCTAAATACTAGG  
TCAGCTTTGGCGACACTGTGTCTTCTCACATAACCACCTGTAGCAAGATGGATCATAAATGAGAAGTGTTCCT  
ATTGATTTAAAGCTTATTGGAATCATGTCTCTTGTCTCTTCGTCCTTTCTTTGCTTTTCTTCTAATTTTCCCTC  
TAGCCTCTCCTCGCCACAATTTGCTGCTTACTGCTGGTGTTAATATTTGTGTGGGATGAATTCCTATCAGGACAA  
CCACTTCTCGAACTGTAATAATGAAGATAATAATATCTTTATCTTTATCCCCCTTCAAGAAATTACCTTTGTG  
TCAAATGCCGCTTTGTTGAGCCCTTAAATACCACCTCCTCATGTGTAAATTGACACAATCACTAATCTGGTAAT  
TTAAACAATTGAGATAGCAAAAGTGTTTAAACAGACTAGGATAATTTTTTTTCATATTTGCCAAAATTTTTGTAA  
ACCCTGTCTTGTCAAATAAGTGATAATATTGTATTATTAATTTATTTTACTTTCTATACCATTTCAAACACA  
TTACACTAAGGGGGAACCAAGACTAGTTTCTTCAGGGCAGTGGACGTAGTAGTTTGTA AAAACGTTTTCTATGAC  
GCATAAGCTAGCATGCCATGATTTATTTCTTCTCATGAATTTGTCACTGGATCAGCAGCTGTGGAAATAAAGCTT  
GTGAGCCCTCTGCTGGCCACAGTGAGGAAAGTAGCACAAATAGGATACAGTTGTATGTAGTCATTGGCAACAATT  
GCATACAATTTTACTACCAAGAGAAGGTATAGTATGGAAAGTCCAATGACTTCCTTGATTGGATGTTAACAGCT  
GACTGGTGTGAGACTTGAGGTTTCATCTAGTCTTCAAACTATATGGTTGCCTAGATTCTCTCTGGAACTGAC  
TTTGTCAAATAAATAGCAGATTGTAGTGTCAAAAAA

263/615

**FIGURE 262**

MDDFISISLLSLAMLVGCYVAGIIPLAVNFSEERLKLVTVLGAGLLCGTALAVIVPEGVHALYEDILEGKHHQAS  
ETHNVIASDKAAEKSVVHEHEHSHDHTQLHAYIGVSLVLGFVFMLLVDQIGNSHVHSTDDPEAARSSNSKITTTTL  
GLVVHAAADGVALGAAASTSQTSLIVFVAIMLHKAPAAFGLVSFLMHAGLERNRIRKHLVLFALAAPVMSMVT  
YLGLSKSSKEALSEVNATGVAMLFSAFTFLYVATVHVLPEVGGIGHSHKPDATGGRGLSRLEVAALVLGCLIPLI  
LSVGHQH

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 37-56, 106-122, 211-230, 240-260, 288-304

264/615

**FIGURE 263**

CTCCTTAGGTGGAAACCCCTGGGAGTAGAGTACTGACAGCAAGACCCGGGAAAGACCATACGTCCCCGG  
GCAGGGGTGACAAACAGGTGTCATCTTTTTGATCTCGTGTGTGGCTGCCTTCTATTTCAAGGAAAGAC  
GCCAAGGTAATTTTGACCCAGAGGAGCAATGATGTAGCCACCTCCTAACCTTCCCTTCTTGAACCCCC  
AGTTATGCCAGGATTTACTAGAGAGTGTCAACTCAACCAGCAAGCGGCTCCTTCGGCTTAACCTTGTGG  
TTGGAGGAGAGAACCTTTGTGGGGCTGCGTTCTCTTAGCAGTGCTCAGAAGTGACTTGCCTGAGGGTG  
GACCAGAAGAAAGGAAAGGTCCCCTCTTGCTGTGGCTGCACATCAGGAAGGCTGTGATGGGAATGAA  
GGTGAACAACTTGGAGATTTCACTTCAGTCATTGCTTCTGCCTGCAAGATCATCCTTTAAAGTAGAGA  
AGCTGCTCTGTGTGGTGGTTAACTCCAAGAGGCAGAACTCGTTCTAGAAGGAAATGGATGCAAGCAGC  
TCCGGGGGGCCCCAAACGCATGCTTCCTGTGGTCTAGCCCAGGGAAGCCCTTCCGTGGGGGGCCCCGGCT  
TTGAGGGATGCCACCGGTTCTGGACGCATGGCTGATTCTGAATGATGATGGTTCGCCGGGGGGCTGCT  
TGGCTGGATTTCCCGGGTGGTGGTTTTGCTGGTGTCTCCTCTGCTGTGCTATCTCTGTCTGTACATGT  
TGGCCTGCACCCCAAAGGTGACGAGGAGCAGCTGGCACTGCCCAGGGCCAAACAGCCCCACGGGGAAG  
GAGGGGTACCAGGCCGTCTTCAGGAGTGGGAGGAGCAGCACCAGCAACTACGTGAGCAGCCTGAAAGC  
GCAGATCGACAGCTCAAGGAGGAGCTGCAGGAGGAGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGC  
CCAGCGATGCTGCTGGCCTGGGTCTGGACAGGAGCCCCCAGAGAAAACCCAGGGCCGACCTCCTGGCC  
TTCCTGCACTCGCAGGTGGACAAGGCAGAGGTGAATGCTGGCGTCAAGCTGGCCACAGAGTATGCAGC  
AGTGCCTTTTCGATAGCTTTACTCTACAGAAGGTGTACCAGCTGGAGACTGGCCTTACCCGCCACCCCG  
AGAGAAGCCTGTGAGGAAGGACAAGCGGGATGAGTTGGTGGAAAGCCATTGAATCAGCGAGTTGAGACC  
CTGAACAATCCTGCAGAGAACAGCCCCAATCACCGTCTTACACGGCCTCTGATTTTCATAGAAGGGAT  
CTACCGAACAGAAAGGGACAAAGGGACATTGTATGAGCTCACCTTCAAAGGGGACCACAAACACGAAT  
TCAAACGGCTCATCTTATTTTCGACCACTTACGCCCCATCATGAAAGTGAAAAATGAAAAGCTCAACATG  
GCCAACACGCTTATCAATGTTATCGTGCTCTAGCAAAAAGGGTGGACAAGTTCGGCAGTTTCATGCA  
GAATTTTCAGGGAGATGTGCATTGAGCAGGATGGGAGAGTCCATCTCACTGTTGTTTACTTTGGGAAAG  
AAGAAATAAATGAAGTCAAAGGAATACTTGAAAACACTTCCAAAGCTGCCAACTTCAGGAACCTTTACC  
TTCATCCAGCTGAATGGAGAATTTTCTCGGGGAAAGGGACTTGATGTTGGAGCCCGCTTCTGGAAGGG  
AAGCAACGCTCTTCTCTTTTCTGTGATGTGGAGCTACTTTCACATCTGAATTCCTCAATACGTGTA  
GGCTGAATACACAGCCAGGGAAGAAGGTATTTTATCCAGTTCTTTTCAGTCAGTACAATCCTGGCATA  
ATATACGGCCACCATGATGCAGTCCCTCCCTTGGAAACAGCAGCTGGTCATAAAGAAGGAAACTGGATT  
TTGGAGAGACTTTGGATTGGGATGACGTGTGATGCTCGGTGAGCTTCAATATAGGTGGGTTTG  
ATCTGGACATCAAAGGCTGGGGCGGAGAGGATGTGCACCTTTATCGCAAGTATCTCCACAGCAACCTC  
ATAGTGGTACGGACGCCTGTGCGAGGACTCTTCCACCTCTGGCATGAGAAGCGCTGCATGGACGAGCT  
GACCCCCGAGCAGTACAAGATGTGCATGCAGTCCAAGGCCATGAACGAGGCATCCCACGGCCAGCTGG  
GCATGCTGGTGTTCAGGCACGAGATAGAGGCTCACCTTCGCAAAACAGAAACAGAACAGTAGAGCAAA  
AAAACATGAACCTCCAGAGAAGGATTGTGGGAGACACTTTTTCTTTCTTTTGAATTAAGTGAAGTG  
GCTGCAACAGAGAAAAGACTTCCATAAAGGACGACAAAAGAATTGGACTGATGGGTGAGAGATGAGAA  
AGCCTCCGATTTCTCTCTGTGGGCTTTTTTACAACAGAAATCAAATCTCCGCTTTGCTGCAAAAGT  
AACCAGTTGCACCCCTGTGAAGTGTCTGACAAAGGCAGAAATGCTTGTGAGATTATAAGCCTAATGGTG  
TGGAGGTTTTGATGGTGTTTACAATACACTGAGACCTGTTGTTTTGTGTGCTCATTGAAATATTCATG  
ATTTAAGAGCAGTTTTGTAAAAAATTCATTAGCATGAAAGGCAAGCATATTTCTCCTCATATGAATGA  
GCCTATCAGCAGGGCTCTAGTTTCTAGGAATGCTAAAATATCAGAAGGCAGGAGAGGAGATAGGCTTA  
TTATGATACTAGTGAGTACATTAAGTAAATAAAATGGACCAGAAAAGAAAACCATAAATATCG  
TGTCATATTTTCCCCAAGATTAAACCAAAAATAATCTGCTTATCTTTTTGGTTGCTTTTTAACTGTCT  
CCGTTTTTTTTCTTTTATTTAAAAATGCACTTTTTTTCCCTTGTGAGTTATAGTCTGCTTATTTAATTA  
CCACTTTGCAAGCCTTACAAGAGAGCACAAGTTGGCCTACATTTTTTATATTTTTTAAAGAGATACTTT  
GAGATGCATTATGAGAACTTTCAAGTTCAAAGCATCAAATTTGATGCCATATCCAAGGACATGCCAAATG  
CTGATTCGTGCAGGCACTGAATGTCAGGCATTGAGACATAGGGAAGGAATGGTTTGTACTAATACAGA  
CGTACAGATACTTTCTCTGAAGAGTATTTTCGAAGAGGAGCAACTGAACACTGGAGGAAAAGAAAATG  
ACACTTTCTGCTTTACAGAAAAGGAAACTCATTCAGACTGGTGATATCGTGATGTACCTAAAAGTCAG  
AAACCACATTTTCTCCTCAGAAGTAGGGACCGCTTCTTACCTGTTTAAATAAACCAAGTATACCGT  
GTGAACCAACAATCTCTTTTCAAACAGGGTGCTCCTCCTGGCTTCTGGCTTCCATAAGAAGAAATG  
GAGAAAATATATATATATATATATATATATTGTGAAAGATCAATCCATCTGCCAGAATCTAGTGGGATG  
GAAGTTTTTGTCTACATGTTATCCACCCAGGCCAGGTGGAAGTAACTGAATTATTTTTTAAATTAAGC  
AGTCTCAATTATCACCAAGATGCTTCTGAAGAAATGCAATTTTATTACCATTTCAACTTTTTTAA  
AATAAATACAGTTAACATAGAGTGGTTTTCTTCATTCATGTGAAAATTATTAGCCAGCACCAGATGCAT  
GAGCTAATTATCTCTTTGAGTCTTGTCTTGTGTTGCTCACAGTAAACTCATTGTTTAAAGCTTCAA  
GAACATTCAAGCTGTTGGTGTGTTAAAAATGCATTGTATTGATTTGTACTGGTAGTTTATGAAATTT  
AATTAACACAGGCCATGAATGGAAGGTGGTATTGCACAGCTAATAAAATATGATTTGTGGATATGAA

265/615

**FIGURE 264**

MMVRRGLLAWISRVVLLVLLCCAISVLYMLACTPKGDEEQLALPRANSPTGKEGYQAVLQEWEEQHRNYVSSL  
KRQIAQLKEELQERSEQLRNGOYQASDAAGLGLDRSPPEKTQADLLAFLHSQVDKAEVNAGVKLATEYAAVPFDS  
FTLQKVYQLETGLTRHPEEKPVKDKRDELVEAIESALETNNPAENSPNHRPYTASDFIEGIYRTERDKGTLYE  
LTFKGDHKHEFKRLILFRPFSPIMKVKNELNMANTLINVIVPLAKRVDKFRQFMQNFREMCIEQDGRVHLTVVY  
FGKEEINEVKGILENTSKAANFRNFTFIQLNGEFSRGKGLDVGARFWKGSNVLLFFCDVDIYFTSEFLNLCRLNT  
QPGKKVFYPVLFSEQYNPGIYGHHDVPPLEQQLVIKKETGFWRDFGFGMTCQYRSDFINIGGFDDIKGWGGED  
VHLYRKYLHSNLIVVRTPVRLFHLWHEKRCMDELTPEQYKCMQSKAMNEASHGQLGMLVFRHEIEAHLRKQKQ  
KTSSKKT

266/615

**FIGURE 265**

GGATGCAGAAAGCCTCAGTGTGCTCTTCCTGGCCTGGGTCTGCTTCCTCTTCTACGCTGGCATTGCCCTCTTCA  
CCAGTGGCTTCCTGCTCACCCTTTGGAGCTCACCAACCATAGCAGCTGCCAAGAGCCCCAGGCCCTGGGTCCC  
TGCCATGGGGGAGCCAAGGGAACCTGGGGCCTGCTGGATGGCTTCCGATTTTCGCGGGTGTGTGGTGTCTGA  
TAGATGCTCTGCGATTTGACTTCGCCCAGCCCCAGCATTCACACGTGCCTAGAGAGCCTCCTGTCTCCCTACCCT  
TCCTGGGCAACTAAGCTCCTTGACAGAGGATCCTGGAGATTGAGCCCCACCATGCCCGGCTCTACCGATCTCAGG  
TTGACCCTCCTACCACCACCATGCAGCGCTCAAGGCCCTCACCCTGGCTCACTGCCACCTTTATTGATGCTG  
GTAGTAACCTCGCCAGCCACGCCATAGTGAAGACAATCTCATTAAAGCAGCTCACCAGTGCAGGAAGCGTGTAG  
TCTTCATGGGAGATGATACCTGGAAAGACCTTTTCCCTGGTGTCTTCTCCAAAGCTTTCTTCTTCCCATCCTTCA  
ATGTCAGAGACCTAGACACAGTGGACAATGGCATCCTGGAACACCTCTACCCACCATGGACAGTGGTGAATGGG  
ACGTGCTGATTGCTCACTTCCTGGGTGTGGACCACTGTGGCCACAAGCATGGCCCTCACCACCCTGAAATGGCCA  
AGAACTTAGCCAGATGGACCAGGTGATCCAGGGACTTGTGGAGCGTCTGGAGAATGACACACTGCTGGTAGTGG  
CTGGGGACCATGGGATGACCACAAATGGAGACCATGGAGGGGACAGTGAGCTGGAGGTCTCAGCTGCTCTCTTTC  
TGTATAGCCCCACAGCAGTCTTCCCAGCACCCCCACCAGAGGAGCCAGAGGTGATTCTCAAGTTAGCCTTGTGC  
CCACGCTGGCCCTGCTGCTGGGCCTGCCCATCCCATTGGGAATATCGGGGAAGTGATGGCTGAGCTATTCTCAG  
GGGTGAGGACTCCCAGCCCCACTCCTCTGCTTTAGCCCCAGCCTCAGCTCTCCATCTCAATGCTCAGCAGGTGT  
CCCGATTTCTCATACCTACTCAGCTGCTACTCAGGACCTTCAAGCTAAGGAGCTTCATCAGCTGCAGAACCTCT  
TCTCCAAGGCCCTCTGCTGACTACCACTGGCTTCTCCAGAGCCCCAAGGGGGCTGAGGCGACACTGCCGACTGTGA  
TTGCTGAGCTGCAGCAGTTCCTGCGGGGAGCTCGGGCCATGTGCATCGAGTCTTGGGCTCGTTCTCTCTGGTCC  
GCATGGCGGGGGTACTGCTCTCTGGCTGCTTCCTGCTTTATCTGCCTGCTGGCATCTCAGTGGGCAATATCCC  
CAGGCTTTCCATCTGCCCTCTACTCCTGACACCTGTGGCCTGGGGCCTGGTTGGGGCCATAGCGTATGCTGGAC  
TCCTGGGAACATTAGAGCTGAAGCTAGATCTAGTGCTTCTAGGGGCTGTGGCTGCAGTGAGCTCATTCTCCCTT  
TTCTGTGGAAGCCTGGGCTGGCTGGGGTCCAAGAGGCCCTGGCAACCCTGTTTCCATCCCTGGGCCCCCTCC  
TGTTACTCCTGCTGTTTTGCTGGCTGTGTTCTTCTCTGATAGTTTTGTTGTAGCTGAGGCCAGGGCCACCCCT  
TCCTTTTGGGCTCATTATCCTGCTCCTGGTGTGCCAGCTTCACTGGGAGGGCCAGCTGCTTCCACCTAAGCTAC  
TCACAATGCCCCGCTTGGCACTTCAGCCACAACAACCCCCACGGCACAATGGTGCATATGCCCTGAGGCTTG  
GAATTGGGTGCTTTTATGTACAAGGCTAGCTGGGCTTTTTCATCGTTGCCCTGAAGAGACACCTGTTGCCACT  
CCTCTCCCTGGCTGAGTCTCTGGCATCCATGGTGGGTGGTTCGAGCCAAGAATTTATGGTATGGAGCTTGTGTGG  
CGGCGCTGGTGGCCCTGTTAGCTGCCGTGCGCTTGTGGCTTCGCGCTATGGTAATCTCAAGAGCCCCGAGCCAC  
CCATGCTCTTTGTGCGCTGGGACTGCCCTAATGGCATTGGGTACTGCTGCTACTGGGCATTTGGCGTGGGGG  
CAGATGAGGCTCCCCCGCTCTCCGGGTCTGGTCTCTGGGGCATCCATGGTGTGCTCGGGCTGTAGCAGGGC  
TGGCTGCTTCAGGGCTCGCGCTGCTGCTCTGGAAGCCTGTGACAGTGTGGTGAAGGCTGGGGCAGGCGCTCCAA  
GGACCAGGACTGTCTCACTCCCTTCTCAGGCCCCCCACTTCTCAAGCTGACTTGGATTATGTGGTCCCTCAAA  
TCTACCGACACATGCAGGAGGAGTTCCGGGGCCGGTTAGAGAGGACCAAATCTCAGGGTCCCCTGACTGTGGCTG  
CTTATCAGTTGGGGAGTGTCTACTCAGCTGCTATGGTCACAGCCCTCACCCTGTTGGCCTTCCACTTCTGCTGT  
TGCATGCGGAGCGCATCAGCCTTGTGTTCCCTGCTTCTGTTTCTGCAGAGCTTCTTCTCCTACATCTGCTTGTG  
CTGGGATACCCGTCAACACCCCTGGTCTTTTACTGTGCCATGGCAGGCAGTCTCGGCTTGGGCCCTCATGGCCA  
CACAGACCTTCTACTCCACAGGCCACCAGCCTGTCTTCCAGCCATCCATTGGCATGCAGCCTTCTGGGATTCC  
CAGAGGGTCAATGGCTCCTGTACTTGGCTGCCTGCTTGTAGTGGGAGCCAACACCTTGGCTCCCACCTCCTCT  
TTGCAGTAGGTTGCCCACTGCTCCTGCTCTGGCCTTCTCTGTGTGAGAGTCAAGGGCTGCGGAAGAGACAGCAGC  
CCCCAGGGAATGAAGCTGATGCCAGAGTCAGACCCGAGGAGGAAGAGGAGCCACTGATGGAGATGCGGCTCCGGG  
ATGCGCCTCAGCACTTCTATGCAGCACTGCTGCAGCTGGGCCCTCAAGTACCTCTTTATCCTTGGTATTCAGATTC  
TGGCCTGTGCTTGGCAGCCTCCATCCTTCGCAGGCATCTCATGGTCTGGAAAGTGTGGCCCTAAGTTCATAT  
TTGAGGCTGTGGGCTTCATTGTGAGCAGCGTGGGACTTCTCCTGGGCATAGCTTTGGTGTGAGAGTGGATGGTG  
CTGTGAGCTCCTGGTTCAGGCAGCTATTTCTGGCCAGCAGAGGTAGCCTAGTCTGTGATTACTGGCACTTGGCT  
ACAGAGAGTGTGGAGAACAGTGTAGCCTGGCCTGTACAGGTACTGGATGATCTGCAAGACAGGCTCAGCCATAC  
TCTTACTATCATGCAGCCAGGGGCCGCTGACATCTAGGACTTCATTATTCTATAATTAGGACCACAGTGGAGTA  
TGATCCCTAACTCCTGATTTGGATGCATCTGAGGACAAGGGGGCGGTCTCCGAAGTGGAATAAATAGGCCGG  
GCGTGGTGACTTGCACCTATAATCCAGCACTTTGGGAGGCAGAGGTGGGAGGATTGCTGGTCCCAGGAGTTCA  
AGACCAGCCTGTGGAACATAACAAGACCCCGTCTACTATTTAAAAAAGTGAATAAATGATAATAT

267/615

**FIGURE 266**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62809

&lt;subunit 1 of 1, 1089 aa, 1 stop

&lt;MW: 118699, pI: 8.49, NX(S/T): 2

MQKASVLLFLAWVCFIFYAGIALFTSGFLLTRLELTNHSSCQEPFPGSLPWGSQGKPGACWMASRFSRVVLVLI  
DALRFDFAPQHSHPREPPVSLPFLGKLSSLQRILEIQPHHARLYRSQVDPPTTTMQRLKALTGSLPTFIDAG  
SNFASHAIVEDNLIKQLTSAGRRVVFMDGDTWKDLFGAFSKAFFPFSFNVRDLDTVDNGILEHLYPTMDSGEWD  
VLIAHFLGVHDHCGHKHGHHPHPEMAKKLSQMDQVIQGLVERLENDTLLVVAGDHGTTNGDHGGDSELEVSAALFL  
YSPTAVFPSTPPEEPEVIPQVSLVPTLALLGLPIPFNGNIGEVMAELFSGGEDSQPHSSALAAQASALHLNAQQVS  
RFLHTYSAATQDLQAKELHQLQNLFSKASADYQWLLQSPKGAETLPTVIAELQQFLRGARAMCIESWARFSLVR  
MAGGTALLAASCFCICLLASQWAI SPGFPPCPLLLTPVAWGLVGAIAYAGLLGTIELKLDLVLLGAVAAVSSFLPF  
LWKAWAGWGSKRPLATLFPPIPGPVLLLLLLFRLAVFFSDSFVVAEARATPFLLGSFILLLVVQLHWEGQLLPKLL  
TMPRLGTSATTNPPRHNGAYALRLGIGLLLCTRLAGLFHRCPEETPVCHSSPWLSPLASMVGGRAKNLWYGACVA  
ALVALLAAVRLWRLRYGNLKSPEPPMLFVRWGLPLMALGTAAYWALASGADEAPPRLRVLVSGASMVLPRAVAGL  
AASGLALLLWKPVTVLVKAGAGAPRTTRTVLTPFSGPPTSQADLDYVVPQIYRHMQEEFRGRLETKSQGPLTVAA  
YQLGSVYSAMVTALTLLAFPLLLHAERISLVFLLFLQSFLLHLLAAGIPVTTPGPFTVPWQAVSAWALMAT  
QTFYSTGHQPVFPPIHWHAAFVGFPEGHGSCTWLPALLVGANTFASHLLFAVGCPLLLLPFLCESQGLRKRQQP  
PGNEADARVRPEEEEEPIMEMRLRDAPQH FYAALLQLGLKYLFI LGIQLACALAASILRRHLMVWKVFAPKFI F  
EAVGFIVSSVGLLLGIALVMRVDGAVSSWFRQLFLAQQR

**Important features:****Signal peptide:**

amino acids 1-16

**Transmembrane domains:**amino acids 317-341, 451-470, 481-500, 510-527, 538-555, 831-850, 1016-1034,  
1052-1070**Leucine zipper pattern.**

amino acids 843-864

**N-glycosylation sites.**

amino acids 37-40, 268-271

**FIGURE 267**

[illegible]

269/615

**FIGURE 268**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62812
><subunit 1 of 1, 337 aa, 1 stop
><MW: 37668, pI: 6.27, NX(S/T): 1
MLFSALLLEVIWILAADGGQHWTYEGPHGQDHWFPASYPECGNNAQSPIDIQTDSVTFDPLPALQPHGYDQPGTE
PLDLHNNNGHTVQLSLPSTLYLGGLPRKYVAAQLHLHWGQKSPGGSEHQINSEATFAELHIVHYDSDSYDSLSEA
AERPQGLAVLGILIEVGETKNIAYEHILSHLHEVRHKDQKTSVPPFNLRELLPKQLGQYFRYNGSLTTPPCYQSV
LWTVFYRRSQISMEQLEKLGTLFSTEEPSKLLVQNYRALQPLNQRMVVFASFIQAGSSYTTGEMLSLGVGILVG
CLCLLLAVYFIARKIRKKRLENRKS VVFTSAQATTEA
```

**Important features of the protein:****Signal peptide:**

amino acids 1-15

**Transmembrane domain:**

amino acids 291-310

**N-glycosylation site.**

amino acids 213-216

**Eukaryotic-type carbonic anhydrases proteins**

amino acids 197-245, 104-140, 22-69



270/615

**FIGURE 269**

GTGGCGCTGGCGGTTGCTGTGCTGATTCCCGGGGTTGGTGGCAGCGGCGGTAGCAGCAATGGACTTTCTCCTG  
GGGAACCCGTTGAGCTCTCCAGTGGGACAGCGCATCGAGAAAGCCACAGATGGCTCCCTGCAGAGCGAGGACTGG  
GCCCTCAACATGGAGATCTGCGACATCATCAACGAGACGGAGGAAGGTCCCAAAGATGCCCTCCGAGCAGTAAAG  
AAGAGAATCGTGGGGAATAAGAACTTCCACGAGGTGATGCTGGCTCTCACAGTCTTAGAAACCTGTGTCAAGAAC  
TGCGGGCACCCTTCCACGTGCTGGTGGCCAGCCAGGACTTCGTGGAGAGTGTGCTGGTGGAGGACCATCCTGCCC  
AAGAACAACCCACCCACCATCGTGCATGACAAAGTGCTCAACCTCATCCAGTCCTGGGCTGACGCGTTCGCGAGC  
TCGCCCCGATCTGACAGGTGTGGTCACCATCTATGAGGACCTGCGGAGGAAAGGCCTGGAGTTCCCATGACTGAC  
CTGGACATGCTGTACCCATCCACACACCCAGAGGACCGTGTCAACTCAGAGACACAATCAGGACAGGATTCTG  
TGGGCACTGACTCCAGCCAGCAAGAGGACTCTGGCCAGCATGCTGCCCCTCTGCCCCCCCCGCCATACTCTCCG  
GTGACACGCCCCATAGCACCACCCCGAACAGATTGGGAAGCTGCGCAGTGAGCTGGAGATGGTGAGTGGGAACG  
TGAGGGTGATGTGCGGAGATGCTGACGGAGCTGGTGCCACCCAGGCCGAGCCCCGAGACCTGGAGCTGCTGCAGG  
AGCTCAACCGCACGTGCCGAGCCATGCAGCAGCGGGTCTGAGTGATAACCCTGCTCCGGGCCCATGCCCCAAGGA  
GCCCTTCAGAGCCCACACTGCCAGTCGAGGCCTGGCTGGAGGCTGGCCACAGTGGAAATTCTGCCGAGCCTATTG  
TCCCTACCCTGCTCTGCTGCATGGGGCCCCATGGCTTTGGCTGGCCACTGAGGGTAGGGTGTGGAGGTGTGGAGG  
CCCCCTGAGGAGCTGCGGCGGGCCAGGTACGAAGCTGCAACTCTGCGCGCAGTGGGCGAGATCTCATCAGCCCCA  
GGCTGCAGGTGAGGCTTCAGGGGATGCTGGGGCCCCACTGCCCCCTCCGCTGCCTTGCCCTCCATCCTTCCTCTGT  
TCCTTCTGGCCGGGCACCACAGCACTGGGGCTCACCTCTTGGTTGATCCTCTTGTAAGTGGGAGAGGTGCCTTTTG  
TATCCCCAATTAAAGGTAGAAAACC

271/615

**FIGURE 270**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62813
><subunit 1 of 1, 209 aa, 1 stop
><MW: 23465, pI: 7.57, NX(S/T): 1
MDFLLGNPFSSPVGQRIEKATDGSLSQSEDWALNMEICDIINETEEGPKDALRAVKKRIVGNKNFHEVMLALTVLE
TCVKNCGHRFHVIVASQDFVESVLVRTLILPKNNPPTIVHDKVLNLIQSWADAFRSSPDLTGVVTIYEDLRRKGLE
FPMTDLDMLSPIHTPRGPCSTQRHNQDRILWALTPASKRTLASMLPLCPPRPYSPVTRP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-15

**N-glycosylation site:**

Amino acids 41-45

**N-myristoylation sites:**

Amino acids 6-12;23-29

272/615

**FIGURE 271**

CGGACGCGTGGSCGGACGCGTGGGCGGACGCGTGGGTCTCTGCGGGGAGACGCCAGCCTGCGTCTGCC**ATG**GGGC  
TCGGGTGAGGGGCTGGGGACGTCTCTGCTGACTGTGGCCACCGCCCTGATGCTGCCCCGTGAAGCCCCCGCAG  
GCTCCTGGGGGGCCAGATCATCGGGGGCCACGAGGTGACCCCCCACTCCAGGCCCTACATGGCATCCGTGCGCT  
TCGGGGGCCAACATCACTGCGGAGGCTTCCTGCTGCGAGCCCGCTGGGTGGTCTCGGCCGCCCACTGCTTCAGCC  
ACAGAGACCTCCGCACTGGCCTGGTGGTGTGGGCGCCACGTCTGAGTACTGCGGAGCCCAACCAGCAGGTGT  
TTGGCATCGATGCTCTCACCACGCACCCCGACTACCACCCCATGACCCACGCCAACGACATCTGCCTGCTGCGGC  
TGAACGGCTCTGCTGTCTGGGCCCTGCAGTGGGGCTGCTGAGGCTGCCAGGGAGAAGGGCCAGGCCCCCACAG  
CGGGGACACGGTGCCGGGTGGCTGGCTGGGGCTTCGTGTCTGACTTTGAGGAGCTGCCGCCTGGACTGATGGAGG  
CCAAGGTCCGAGTGCTGGACCCGGACGTCTGCAACAGCTCCTGGAAGGGCCACCTGACACTTACCATGCTCTGCA  
CCCGCAGTGGGGACAGCCACAGACGGGGCTTCTGCTCGGCCGACTCCGGAGGGGCCCTGGTGTGCAGGAACCGGG  
CTCACGGCCTCGTTTCTCTCGGGCCTCTGGTGCGGCGACCCCAAGACCCCGACGTGTACACGCAGGTGTCCG  
CCTTTGTGGCCTGGATCTGGGACGTGGTTCGGCGGAGCAGTCCCCAGCCCGGCCCTGCCTGGGACCACCAGGC  
CCCCAGGAGAAGCCGCC**TGA**GCCACAACCTTGCGGCATGCAAATGAGATGGCCGCTCCAGGCCTGGAATGTTCCG  
TGGCTGGGCCCCACGGGAAGCCTGATGTTCAAGGTTGGGGTGGGACGGGCAGCGGTGGGGCACACCCATTCCACA  
TGCAAAGGGCAGAAGCAAACCCAGTAAAATGTAACTGACAAAAAAAAAAAAAAAAAAGAAA

273/615

**FIGURE 272**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62845
><subunit 1 of 1, 283 aa, 1 stop
><MW: 30350, pI: 9.66, NX(S/T): 2
MGLGLRGWGRPLLTVATALMLPVKPPAGSWGAIIGGHEVTPHSRPYMASVRFGGQHHCGGFLLRARWVVSAAHC
FSHRDLRTGLVVLGAHVLSTAEP TQQVFGIDALTTHPDYHPMTHANDICLLRLNGSAVLGPAVGLLRLPGRRARP
PTAGTRCRVAGWGFVSDFEELPPGLMEAKVRVLDPDVCNSSWKGHLLTLMCLTRSGDSHRRGFCSADSGGPLVCR
NRAHGLVSFSGLWCGDPKTPDVYTQVSAFVAWIWDVVRSSPQPGPLPGTTRPPGEAA
```

**Signal peptide:**  
amino acids 1-30

274/615

**FIGURE 273**

GAAGTTCGCGAGCGCTGGCT**ATG**GGTCCTGGGGCGCGGCTGGCGGGCGCTGCTGGCGGTGCTGGCGCTCGGGACAG  
GAGACCCAGAAAGGGCTGCGGCTCGGGGCGACACGTTCTCGGCGCTGACCAGCGTGGCGCGCGCCCTGGCGCCCC  
AGCGCCGGCTGCTGGGGCTGCTGAGGCGGTACCTGCGCGGGGAGGAGGCGCGGCTGCGGGACCTGACTAGATTCT  
ACGACAAGGTACTTTCTTTGCATGAGGATTCAACAACCCCTGTGGCTAACCCCTCTGCTTGCATTTACTCTCATCA  
AACGCCTGCAGTCTGACTGGAGGAATGTGGTACATAGTCTGGAGGCCAGTGAGAACATCCGAGCTCTGAAGGATG  
GCTATGAGAAGGTGGAGCAAGACCTTCCAGCCTTTGAGGACCTTGAGGGAGCAGCAAGGGCCCTGATGCGGCTGC  
AGGACGTGTACATGCTCAATGTGAAAGGCCTGGCCCGAGGTGTCTTTAGAGAGTCACTGGCTCTGCCATCACTG  
ACCTGTACAGCCCCAAACGGCTCTTTCTCTCACAGGGGATGACTGCTTCCAAGTTGGCAAGGTGGCCTATGACA  
TGGGGGATTATTACCATGCCATTCCATGGCTGGAGGAGGCTGTCAGTCTCTTCCGAGGATCTTACGGAGAGTGGA  
AGACAGAGGATGAGGCAAGTCTAGAAGATGCCTTGGATCACTTGGCCTTTGCTTATTTCCGGGCAGGAAATGTTT  
CGTGTGCCCTCAGCCTCTCTCGGGAGTTTCTTCTCTACAGCCCAGATAATAAGAGGATGGCCAGGAATGTCTTGA  
AATATGAAAGGCTCTTGGCAGAGAGCCCCAACCACGTGGTAGCTGAGGCTGTATCCAGAGGCCCAATATACCCC  
ACCTGCAGACCAGAGACACCTACGAGGGGCTATGTGAGACCTGGGTCCCAGCCCCTCTCTACAGATCCCTA  
GCCTCTACTGTTCTTATGAGACCAATTCCAACGCCTACCTGCTGCTCCAGCCCATCCGGAAGGAGGTCATCCACC  
TGGAGCCCTACATTGCTCTCTACCATGACTTCGTGAGTCACTGAGAGGCTCAGAAAATTAGAGAACTTGCAGAAC  
CATGGCTACAGAGGTCAGTGGTGGCATCAGGGGAGAAGCAGTTACAAGTGGAGTACCGCATCAGCAAAAGTGCCT  
GGCTGAAGGACACTGTTGACCCAAAAGTGGTGACCTCAACCACCGCATTGCTGCCCTCACAGGCCTTGATGTCC  
GGCCTCCCTATGCAGAGTATCTGCAGGTGGTGAAGTATGGCATCGGAGGACACTATGAGCCTCACTTTGACCATG  
CTACGTACCAAGCAGCCCCCTCTACAGAATGAAGTCAGGAAACCGAGTTGCAACATTTATGATCTATCTGAGCT  
CGGTGGAAGCTGGAGGAGCCACAGCCTTCATCTATGCCAACCTCAGCGTGCTGTGGTTAGGAATGCAGCACTGT  
TTTGGTGGAACTGCACAGGAGTGGTGAAGGGGACAGTGACACACTTCATGCTGGCTGTCTGTCTGCTGGGAG  
ATAAGTGGGTGGCCAACAAGTGGATAACATGAGTATGGACAGGAATTCCGCAGACCCTGCAGCTCCAGCCCTGAAG  
ACT**GAA**ACTGTTGGCAGAGAGAAGCTGGTGGAGTCTGTGGCTTCCAGAGAAGCCAGGAGCCAAAAGCTGGGGTA  
GGAGAGGAGAAAGCAGAGCAGCCTCCTGGAAGAAGGCCTTGTGAGCTTGTCTGTGCCTCGCAAATCAGAGGCAA  
GGGAGAGGTTGTTACAGGGGACACTGAGAATGTACATTTGATCTGCCCCAGCCACGGAAGTCAGAGTAGGATGC  
ACAGTACAAAGGAGGGGGGAGTGGAGGCCTGAGAGGGAAGTTTCTGGAGTTCAGATACTCTCTGTTGGGAACAGG  
ACATCTCAACAGTCTCAGGTTGATCAGTGGGTCTTTTGGCACTTTGAACCTTGACCACAGGGACCAAGAAGTGG  
CAATGAGGACACCTGCAGGAGGGGCTAGCCTGACTCCCAGAACTTTAAGACTTTCTCCCCACTGCCTTCTGCTGC  
AGCCCCAAGCAGGGAGTGTCCCCCTCCCAGAAGCATATCCAGATGAGTGGTACATTATATAAGGATTTTTTTTAA  
GTTGAAAACAACCTTTCTTTTCTTTTGTATGATGGTTTTTTAACACAGTCATTAAAAATGTTTATAAATCAAAA

275/615

**FIGURE 274**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64849
><subunit 1 of 1, 544 aa, 1 stop
><MW: 61126, pI: 6.40, NX(S/T): 2
MGPGARLAALLAVLALGTGDPERAAARGDTFSALTSVARALAPERRLGLLRRYLRGEEA
RLRDLTRFYDKVLSLHEDSTTPVANPLLAFTLIKRLQSDWRNVVHSLEASENIRALKDGY
EKVEQDLPAFEDLEGAARALMRLQDVYMLNVKGLARGVFQRVITGSAITDLYSPKRLFSIT
GDDCFQVGKVAIDMGDYYHAIPWLEEAVSLFRGSYGEWKTEDEASLEDALDHLAFAYFRA
GNVSCALSLSREFLLYSPDNKRMAENVLYKYLAEVSPNHVVAEAVIQRPNIPLQTRDT
YEGLCQTLGSOPTLYQIPSLYCSYETNSNAYLLLQPIRKEVIHLEPYIALYHDFVSDSEA
QKIRELAEPWLQRSVVASGEKQLQVEYRISKSAWLKDTVDPKLVTLNHRIAALTGLDVRP
PYAEYLQVVNYGIGGHYEPHFDHATSPSSPLYRMKSGNRVATFMIYLLSSVEAGGATAFIY
ANLSVPVVRNAALFWWNLHRSGEGSDTLHAGCPVLVGDKWVANKWIEHYGQEFRPCSS
SPED
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-19

**Leucine zipper pattern:**

Amino acids 34-56;41-63

**Ribonucleotide reductase small subunit signature:**

Amino acids 340-356

**N-glycosylation sites:**

Amino acids 242-246;482-486

**Cell attachment sequence:**

Amino acids 27-30

**Tyrosine kinase phosphorylation site:**

Amino acids 189-198

**N-myristoylation sites:**Amino acids 4-10;135-141;153-159;164-170;241-247;303-309;309-315;  
457-463;473-479

276/615

**FIGURE 275**

GGCAACATGGCTCAGCAGGCTTGCCCCAGAGCCATGGCAAAGAATGGACTTGTAATTTGCATCCTGGTGATCACC  
TTACTCCTGGACCAGACCAGCCACACATCCAGATTAAGCCAGGAAGCACAGCAAACGTCGAGTGAGAGAC  
AAGGATGGAGATCTGAAGACTCAAATTTGAAAAGCTCTGGACAGAAGTCAATGCCTTGAAGGAAATTCAGCCCTG  
CAGACAGTCTGTCTCCGAGGCACTAAAGTTCACAAGAAATGCTACCTTGCTTCAGAAGGTTTGAAGCATTTCAT  
GAGGCCAATGAAGACTGCATTTCAAAGGAGGAATCCTGGTTATCCCCAGGAACCTCCGACGAAATCAACGCCCTC  
CAAGACTATGGTAAAAGGAGCCTGCCAGGTGTCAATGACTTTTGGCTGGGCATCAATGACATGGTCACGGAAGGC  
AAGTTTGTGACGTCAACGGAATCGCTATCTCCTTCCTCAACTGGGACCGTGCACAGCCTAACGGTGGCAAGCGA  
GAAACTGTGTCTCTCTCCCAATCAGCTCAGGGCAAGTGGAGTGATGAGGCCTGTGCGAGCAGCAAGAGATAC  
ATATGCGAGTTCACCATCCCTAAATAGGTCTTTCTCCAATGTGTCTCCTCCAAGCAAGATTCATCATAACTTATAGG  
TTCATGATCTCTAAGATCAAGTAAAAATCATAATTTTACTTTATTAATAAATTTGCAACACAAGATCAATGTCCAT  
AGCAATATGATAGCATCAGCCAATTTTGCTAACACATTTCTTTGGGATTTTGCCCTTCTGGGGTATAGGGGATC  
AGAAATATTGATCCATGTGCACGCAGATAAAATGGCTTCTGCTAAACAGACTAAAATCTTTCTCTCTAGTCTTTC  
TCACTTGTAACAAACCCAGTTTGTGTTTTCAAAAAATCACAGTAGCAATGCAACTCATCACTCTAGAAAAGCAAGCTT  
AGGCTACCTGAAAGATTTTCCCTTGGAAGTTTAGCGTATGTTTGACTAACAAAAATTCCTACATCAGAGACTCT  
AGGTGCTATATAATCCAAAACTTTTCAGCCTGTTGCTCATTCTGTCCCATGCTGGCAATAATACCTTGTCAGCC  
CATTACCCTTATTTTGAATTGCTCCATCTCCTGGTGGGACTTGATCTTGCTGCTGCCATATCAGAACACAAACCCC  
TGAAGAGGTTCTGATTTGATTTTTTTTTTTTCTTCATGCCTACCCTTTTTTTTGAAGTTTCCAGCCGCAATTTGA  
AATGAAATGACAAGGTGTATATTTGATCAATTTTCATTCCACCATTGCATTACAACCTCTAACTTAAATGGGTA  
ACCCTAAGGCATATCAAAGAAGCAGATTGCATGATAACGGAATAGAAAAAAGAACCTACATTTATTTTGCTT  
TAGCATCCTTACTCTCACCTTTTATGAGATTGAGAGTGGACTTACATTTCTTTTACATTTTCGTATATTTAT  
TTTTTTTAGCCATCATTATATGTTTAAGTCTATTATGGGCAACCAATCTTTGGAAGCTGAAAAGTGAATTTAAAG  
AATGCTATCTTGAAAAATTGCATACGTCTGTGCAATTTTTTATTCTGCCTAGTGCTATTCTGCTTGTTTAACTAG  
ATTGTACAAAATAACTTCATTGCTTAATATCAAATTACAAAGTTTAGACTTGGAGGGAAATGGGCTTTTTAGAAG  
CAAACAATTTTAAATATATTTTGTCTTCAAATAAATAGTGTTTAAACATTGAATGTGTTTTGTGAACAATATCC  
CACTTTGCAAACTTTAACTACACATGCTTGAATTAAGTTTTAGCTGTTTTTCATTGCTCAATAATAAAGCCTGAA  
TTCTGATCAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

277/615

**FIGURE 276**

MAQQACPRAMAKNGLVICILVITLLLDQTTSHTSRLKARKHRSKRRVRDKDGLKTQIEKLWTEVNALKEIQALQT  
VCLRGTKVHKCYLASEGLKHFHEANEDCISKGGILVIPRNSDEINALQDYGKRSLPGVNDFWLGINDMVTEGKF  
VDVNGIAISFLNWDRAQPNGGKRENCVLFSSQAQKWSDEACRSSKRYICEFTIPK



278/615

**FIGURE 277**

GAGATAGGGAGTCTGGGTTTAAGTTCCTGCTCCATCTCAGGAGCCCCTGCTCCCACCCCTAGGAAGCCACCAGAC  
TCCACGGTGTGGGGCCAATCAGGTGGAATCGGCCCTGGCAGGTGGGGCCACGAGCGCTGGCTGAGGGACCGAGCC  
GGAGAGCCCCGGAGCCCCCGTAACCCGCGCGGGGAGCGCCAGGATGCGCGCGGGGACTCGGAGCAGGTGCGCT  
ACTGCGCGCGCTTCTCCTACCTCTGGCTCAAGTTTTCACTTATCATCTATTCCACCGTGTCTGGCTGATTGGGG  
CCCTGGTCTGTCTGTGGGCATCTATGCAGAGGTTGAGCGGCAGAAATATAAAACCCTTGAAAGTGCCTTCCTGG  
CTCCAGCCATCATCCTCATCCTCCTGGGCGTCGTCATGTTTCATGGTCTCCTTCATTGGTGTGCTGGCGTCCCTCC  
GTGACAACTGTACCTTCTCCAAGCATTTCATGTACATCCTTGGGATCTGCCTCATCATGGAGCTCATTGGTGGCG  
TGGTGGCCTTGACCTTCCGGAACCAGACCATTGACTTCCTGAACGACAACATTGGAAGAGGAATTGAGAACTACT  
ATGATGATCTGGACTTCAAAAACATCATGGACTTTGTTTCAAAAAAGTTCAAGTGCTGTGGCGGGGAGGACTACC  
GAGATTGGAGCAAGAATCAGTACCACGACTGCAGTGCCCTGGACCCCTGGCCTGTGGGGTGCCCTACACCTGCT  
GCATCAGGAACACGACAGAAGTTGTCAACACCATGTGTGGCTACAAAATATCGACAAGGAGCGTTTCAGTGTGC  
AGGATGTCTCTACGTGCGGGGCTGCACCAACGCCGTGATCATCTGGTTCATGGACAACCTACACCATCATGGCGT  
GCATCCTCCTGGGCATCCTGCTTCCCCAGTTTCTGGGGGTGCTGCTGACGCTGCTGTACATCACC CGGTGGAGG  
ACATCATCATGGAGCACTCTGTCACTGATGGGCTCCTGGGGCCCGGTGCCAAGCCAGCGTGGAGGCGGCAGGCA  
CGGGATGCTGCTTGTGCTACCCCAATTAGGGGCCAGCCTGCCATGGCAGCTCCAACAAGGACCGTCTGGGATAGC  
ACCTCTCAGTCAACATCGTGGGGCTGGACAGGGCTGCGGCCCCCTGTGCCACACTCAGTACTGACCAAAGCCAGG  
GCTGTGTGTGCCTGTGTGTAGTCCCACGGCCTCTGCCTCCCCAGGGAGCAGAGCCTGGGCCTCCCCTAAGAGGC  
TTTCCCCGAGGCAGCTCTGGAATCTGTGCCACCTGGGGCCTGGGGAAACAAGGCCCTCCTTTCTCCAGGCCTGGG  
CTACAGGGGAGGGAGAGCCTGAGGCTCTGCTCAGGGCCATTTTCATCTCTGGCAGTGCCTTGGCGGTGGTATTCA  
AGGCAGTTTTGTAGCACCTGTAATTGGGGAGAGGGAGTGTGCCCCTCGGGGCAGGAGGGAAGGGCATCTGGGGAA  
GGGCAGGAGGGAAGAGCTGTCCATGCAGCCACGCCCATGGCCAGGTTGGCCTCTTCTCAGCCTCCCAGGTGCCTT  
GAGCCCTCTTGCAAGGGCGGCTGCTTCTTGGCCTAGTTTTTTTTTACGTGATTTTTGTAACATTCATTTTTTT  
GTACAGATAACAGGAGTTTTCTGACTAATCAAAGCTGGTATTTCCCGCATGTCTTATTCTTGCCCTTCCCCAAC  
CAGTTTGTTAATCAAACAATAAAAACATGTTTTGTTTTGTTTTTAAAAA

279/615

**FIGURE 278**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64863
><subunit 1 of 1, 294 aa, 1 stop
><MW: 33211, pI: 5.35, NX(S/T): 3
MPRGDSEQVRYCARFSYLWLKFSLLIYSTVFWLIGALVLSVGIYAEVERQKYKTLES AFLAPAILILLGVVMFM
VSFIGVLASLRDONLYLLQAFMYILGICLIMELIGGVVALTFRNQTIDFLNDNIRRG IENYYDDLDFKNIMDFVQK
KFKCCGGEDYRDWSKNQYHDCSAPGPLACGVPYTCCIRNTTEVVNTMCGYKTIDKERFSVQDVIYVRGCTNAVII
WFMDNYTIMACILLGILLPQFLGVLLTLLYITRVEDIIMEHSVTDGLLGPGAKPSVEAAGTGCCLCYPN
```

**Signal peptide:**  
amino acids 1-44

**Transmembrane domains:**  
amino acids 22-42, 57-85, 93-116, 230-257

280/615

**FIGURE 279**

GAGGAGCGGGCCGAGGACTCCAGCGTGCCAGGTCTGGCATCCTGCACTTGCTGCCCTCTGACACCTGGGAAGAT  
GGCCGGCCCGTGGACCTTCACCCCTTCTCTGTGGTTTGCTGGCAGCCACCTTGATCCAAGCCACCCCTCAGTCCCAC  
TGCAGTTCTCATCCTCGGCCCAAAGTCATCAAAGAAAAGCTGACACAGGAGCTGAAGGACCACAACGCCACCAG  
CATCCTGCAGCAGCTGCCGCTGCTCAGTGCCATGCGGGAAAAGCCAGCCGGAGGCATCCCTGTGCTGGGCAGCCT  
GGTGAACACCGTCCTGAAGCACATCATCTGGCTGAAGGTCATCACAGCTAACATCCTCCAGCTGCAGGTGAAGCC  
CTCGGCCAATGACCAGGAGCTGCTAGTCAAGATCCCCCTGGACATGGTGGCTGGATTCAACACGCCCCCTGGTCAA  
GACCATCGTGGAGTTCCACATGACGACTGAGGCCCAAGCCACCATCCGCATGGACACCAGTGCAAGTGGCCCCAC  
CCGCTGGTCTCAGTGACTGTGCCACCAGCCATGGGAGCCTGCGCATCCAACCTGCTGTATAAGCTCTCCTTCT  
GGTGAACGCCCTTAGCTAAGCAGGTGATGAACCTCCTAGTGCCATCCCTGCCCAATCTAGTAAAAACCAGCTGTG  
TCCCGTGATCGAGGCTTCCTTCAATGGCATGTATGCAGACCTCCTGCAGCTGGTGAAGGTGCCATTTCCCTCAG  
CATTGACCGTCTGGAGTTTGACCTTCTGTATCCTGCCATCAAGGGTGACACCATTGAGCTCTACCTGGGGGCCAA  
GTTGTTGGACTCACAGGGAAAGGTGACCAAGTGGTTCAATAACTCTGCAGCTTCCCTGACAATGCCACCCTGGA  
CAACATCCCGTTGAGCCTCATCGTGAGTCAGGACGTGGTGAAGCTGCAGTGGCTGCTGTGCTCTCTCCAGAAGA  
ATTCATGGTCTGTTGGACTCTGTGCTTCCCTGAGAGTGCCCATCGGCTGAAGTCAAGCATCGGGCTGATCAATGA  
AAAGGCTGCAGATAAGCTGGGATCTACCCAGATCGTGAAGATCCTAACTCAGGACACTCCCGAGTTTTTATAGA  
CCAAGGCCATGCCAAGGTGGCCCAACTGATCGTGCTGGAAGTGTTCCCTCCAGTGAAGCCCTCCGCCCTTTGTT  
CACCTGGGCATCGAAGCCAGCTCGGAAGCTCAGTTTTACACCAAAGGTGACCAACTTATACTCAACTGAATAA  
CATCAGCTCTGATCGGATCCAGCTGATGAACTCTGGGATTGGCTGGTTCCAACCTGATGTTCTGAAAAACATCAT  
CACTGAGATCATCCACTCCATCCTGCTGCCGAACCAGAATGGCAAATTAAGATCTGGGGTCCCAGTGTCATTGGT  
GAAGGCCTTGGGATTGAGGCAGCTGAGTCCTCACTGACCAAGGATGCCCTTGCTGCTTACTCCAGCCTCCTTGTG  
GAAACCCAGCTCTCCTGTCTCCAGTGAAGACTTGGATGGCAGCCATCAGGGAAGGCTGGGTCCCAGCTGGGAGT  
ATGGGTGTGAGCTCTATAGACCATCCCTCTCTGCAATCAATAAACACTTGCCCTGTGAAAAA

281/615

**FIGURE 280**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64881
><subunit 1 of 1, 484 aa, 1 stop
><MW: 52468, pI: 7.14, NX(S/T): 3
MAGPWTFTLLCGLLAATLIQATLSPTAVLILGPKVIKEKLTQELKDNATSILOQLPLLSAMREKPAGGIPVLGS
LVNTVLKHIIWLKVITANILQLQVKPSANDQELLVKIPLDMVAGFNTPLVKTIVEFHMTTEAQATIRMDTSASGP
TRLVLSDCATSHGSLRIQLLYKLSFLVNALAKQVMNLLVPSLPNLVKNQLCPVIEASFNGMYADLLQLVKVPISL
SIDRLEFDLLYPAIKGDITQLYLGAKL LDSQGKVTKWFNNSAASLTMP TLDNIPFSLIVSQDVVKA AVAVLSPE
EFMVLLDSVLPESAHLKSSIGLINEKAADKLGSTQIVKILTQDTPEFFIDQGHAKVAQLIVLEVFPSSSEALRPL
FTLGIEASSEAQFYTKGDQLILNLNNISSDRIQLMNSGIGWFQPDVLKNIITEIIHSILLPNQNGKLRSGVPVSL
VKALGFEEAAESSLT KDALVLT PASLWKPPSPVSQ
```

**Important features of the protein:****Signal peptide:**

amino acids 1-21

**N-glycosylation sites.**

amino acids 48-51, 264-267, 401-404

**Glycosaminoglycan attachment site.**

amino acids 412-415

**LBP / BPI / CETP family proteins.**

amino acids 407-457

282/615

**FIGURE 281**

CCCACGCGTCCGCGCCTCTCCCTTCTGCTGGACCTTCCTTCGTCTCTCCATCTCTCCCTCCTTTCCCCGCGTTCT  
CTTTCCACCTTTCTCTTCTTCCACCTTAGACCTCCCTTCCTGCCCTCCTTTCCTGCCCCACCGCTGCTTCCTGGC  
CCTTCTCCGACCCCGCTCTAGCAGCAGACCTCCTGGGGTCTGTGGGTTGATCTGTGGCCCCCTGTGCCTCCGTGTC  
CTTTTCGTCTCCCTTCCTCCCGACTCCGCTCCCGGACCAGCGGCCTGACCCTGGGGAAAGGATGGTTCCCCGAGGT  
GAGGGTCTCTCCTCCTTGCTGGGACTCGCGCTGCTCTGGTTCCCCCTGGACTCCACGCTCGAGCCCCGCCAGA  
CATGTTCTGCCTTTTCCATGGGAAGAGATACTCCCCCGCGGAGAGCTGGCACCCCTACTTGGAGCCACAAGGCCT  
GATGTACTGCCTGCGCTGTACCTGCTCAGAGGGCGCCCATGTGAGTTGTTACCGCCTCCACTGTCCGCCTGTCCA  
CTGCCCCCAGCCTGTGACGGAGCCACAGCAATGCTGTCCCAAGTGTGTGGAACCTCACACTCCCTCTGGACTCCG  
GGCCCCACCAAAGTCTGCCAGCACAACGGGACCATGTACCAACACGGAGAGATCTTCAGTGCCCATGAGCTGTT  
CCCCCTCCCGCCTGCCCAACCAAGTGTGTCTCTGCAGCTGCACAGAGGGCCAGATCTACTGCGGCCTCACAACTG  
CCCCGAACCAGGCTGCCCAGCACCCCTCCCACTGCCAGACTCCTGCTGCCAAGCCTGCAAAGATGAGGCAAGTGA  
GCAATCGGATGAAGAGGACAGTGTGCAGTCGCTCCATGGGGTGAGACATCCTCAGGATCCATGTTCCAGTGATGC  
TGGGAGAAAGAGAGGGCCCGGGCACCCAGCCCCCACTGGCCTCAGCGCCCCCTCTGAGCTTCATCCCTCGCCACTT  
CAGACCCAAGGGAGCAGGCAGCACAAGTGTCAAGATCGTCTGAAGGAGAAACATAAGAAAGCCTGTGTGCATGG  
CGGGAAGACGTACTCCCACGGGGAGGTGTGGCACCCGGCCTTCCGTGCCTTCGGCCCCCTTGCCCTGCATCCTATG  
CACCTGTGAGGATGGCCGCCAGGACTGCCAGCGTGTGACCTGTCCCACCGAGTACCCCTGCCGTACCCCCGAGAA  
AGTGGCTGGGAAGTGCTGCAAGATTGCCCAGAGGACAAAGCAGACCCTGGCCACAGTGAGATCAGTTCTACCAG  
GTGTCCCAAGGCACCGGGCCGGGTCTCGTCCACACATCGGTATCCCCAAGCCCAGACAACCTGCGTCTGCTTTGC  
CCTGGAACACGAGGCCTCGGACTTGGTGGAGATCTACCTCTGGAAGCTGGTAAAAGATGAGGAAACTGAGGCTCA  
GAGAGGTGAAGTACCTGGCCCAAGGCCACACAGCCAGAATCTTCCACTTGACTCAGATCAAGAAAGTCAGGAAGC  
AAGACTTCCAGAAAGAGGCACAGCACTTCCGACTGCTCGCTGGCCCCACGAAGGTCACTGGAACGTCTTCTTAG  
CCCAGACCCTGGAGCTGAAGGTCACGGCCAGTCCAGACAAAGTGACCAAGACATAACAAAGACCTAAACAGTTGCA  
GATATGAGCTGTATAATTGTTGTTATTATATATTAATAAATAAGAAGTTGCATTACCCTCAAAAAAAAAAAAAA  
AAAAAA

283/615

**FIGURE 282**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64902
><subunit 1 of 1, 451 aa, 1 stop
><MW: 49675, pI: 7.15, NX(S/T): 1
MVPEVRVLSSLLGLALLWFPLDSHARARPD MFCLFHGKRYSPGESWHPYLEPQGLMYCLRCTCSEGAHVSCYRLH
CPPVHCPQPVTEPQQCCPKC VEPHTPSGLRAPPKSCQHNGTMYQHGEIFSAHELFPSRLPNQCVLCSCTEGQIYC
GLTTCPEPGCPAPLPLPDSCCQACKDEASEQSDEEDSVQSLHGVRHPQDPCSSDAGRKRGP GTPAPTGLSAPLSF
IPRHFRPKGAGSTTVKIVLKEKHKKACVHGKTYSHGEVWHPAFRAFGPLPCILCTCEDGRQDCQRVTCPT EYPC
RHPEKVAGKCKICPEDKADPGHSEISSTRCPKAPGRVLVHTSVSPSPDNLRRFALEHEASDLVEIY LWKLVKDE
ETEAQRGEVPGPRPHSQNLPLDSDQESQEARLPERGTALPTARWP PRRSLERLPSDPGAEGHGQSRQSDQDITKT
```

**Signal peptide:**

amino acids 1-25

284/615

**FIGURE 283**

GCGATGGTGC GCCCGGTGGCGGTGGCGGCGGCGGTTGCGGAGGCTTCCTTGGTCCGATTGCAACGAGGAGAAGAT  
GACTGACCAACCGACTGGCTGAATGAATGAATGGCGGAGCCGAGCGCGCCATGAGGAGCCTGCCGAGCCTGGGCG  
GCCTCGCCCTGTTGTGCTGCGCCGCCGCCGCCGCCGCCGTGCGCTCAGCCGCCCTCGGCGGGGAATGTCACCGGTG  
GCGGCGGGGCGCGGGGCGAGGTGGACGCGTCGCCGGGCCCCGGGTTGCGGGGCGAGCCAGCCACCCCTTCCCTA  
GGGCGACGGCTCCACGGCCAGGCCCCGAGGACCGGGCCCCCGCGCGCCACCGTCCACCGACCCCTGGCTGCGA  
CTTCTCCAGCCAGTCCCCGGAGACCACCCCTCTTTGGGCGACTGCTGGACCCTCTTCCACCACCTTTCAGGCGC  
CGCTCGGCCCTCGCCGACCACCCCTCCGGCGGCGGAACGCACTTCGACCACCTCTCAGGCGCCGACCAGACCCG  
CGCCGACCACCCCTTCGACGACCACTGGCCCGCGCGCCGACCACCCCTGTAGCGACCACTGACCGCGCCACGA  
CTCCCCGACCCCGACCCCGATCTCCCCAGCAGCAGCAACAGCAGCGTCCTCCCCACCCACCTGCCACCGAGG  
CCCCCTCTTCGCTCCTCCAGAGTATGTATGTAAGTCTGTGTTGGTGAAGCCTGAATGTGAATCGCTGCAACC  
AGACCACAGGGCAGTGTGAGTGTGCGCCAGGTTATCAGGGGCTTCACTGTGAAACCTGCAAGAGGGCTTTTACC  
TAAATTACACTTCTGGGCTCTGTGAGCCATGTGACTGTAGTCCACATGGAGCTCTCAGCATACCGTGCAACAGGT  
AAGCAACAGAGGGTGGAACTGAAGTTTATTTTATTTTAGCAAGGGGAAAAAAAAGGCTGCTACTCTCAAGGACCA  
TACTGGTTTAAACAAAGGAGGATGAGGGTCATAGATTTACAAAATATTTTATATACTTTTATTCTCTTACTTTAT  
ATGTTATATTTAATGTCAGGATTTAAAAACATCTAATTTACTGATTTAGTTCTTCAAAGCACTAGAGTCGCCAA  
TTTTTCTCTGGGATAATTTCTGTAAATTTTCATGGGAAAAAATTATTGAAGAATAAATCTGCTTCTGGAAGGGCT  
TTCAGGCATGAAACCTGCTAGGAGGTTTAGAAATGTTCTTATGTTTATTAATATAACATTGGAGTTTGAGGAAAT  
TTGTTGTTTGGTTTATTTTCTCTCTAATCAAATTTCTACATTTGTTTCTTTGGACATCTAAAGCTTAACCTGGG  
GGTACCCTAATTTATTTAACTAGTGGTAAGTAGACTGGTTTACTCTATTTACCAGTACATTTTGGAGACCAAAA  
GTAGATTAAGCAGGAATTATCTTTAACTATTATGTTATTTGGAGGTAATTTAATCTAGTGAATAATGTACTGT  
TATCTAAGCATTTGCCTTGACTGCACTGAAAGTAATTATTCTTTGACCTTATGTGAGGCACCTGGCTTTTTGTG  
GACCCCAAGTCAAAAACTGAAGAGACAGTATTAAATAATGAAAAAATAATGACAGGTTATACTCAGTGTAAAC  
TGGGTATAACCCAGATCTGCTGCCACTTACGAGCTGTGTTCCCTTGGGCAAGTAATTTCTTTCACTGAGCTTGT  
TTCTTCTCAAGGTTGTTGTGAAGATTAAATGAGTTGATATATATAAATGCCTAGCACATGTCACTCAATAAATT  
CTGTTTTGTTTTAATTTCAAAGGAATATTATGGAAGTGAATGAGAGAACATGTTTTAAGAACTTTTAGCTCCTTG  
ACAAAGAAGTGCTTTTATACTTTAGCACTAAATATTTTAAATGCTTTATAAATGATATTATACTGTTATGGAATAT  
TGTATCATATTGTAGTTTATTAATAATGTAGAAGAGGCTGGGCGCGGTGGCTCAGGCCTGTAATCCTAGCACTTT  
GGGAGGCCAAGGCGGTGGATCACTTGAGGCCAGGAGTTCTAGATGAGCCTGGCCAGCACAGTGAAACCCCGTCT  
CTACTAAAAATACAAACAAATTAGCTGGGCGTGGTGGCACACACCTGTAGTCCCAGCTACTCGGGAGGCTGAGGC  
AGGAGAATCGGTTGAACCCGGGAGGTGGAGGTTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGTGAG  
AGAGGGAGACTCTGTCTTAAAAAAAAAAAAAAAAAAAAAAAAA

285/615

**FIGURE 284**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA64952
><subunit 1 of 1, 258 aa, 1 stop
><MW: 25716, pI: 8.13, NX(S/T): 5
MRSLSLGGGLALLCCAAAAAASASAASAGNVTGGGGAAGQVDASPGPGLRGEPSHFFPRATAPTAQAPRTGPPRA
TVHRPLAATSPAQSPETTPWATAGPSSTTFQAPLGPSPTTPPAERTSTTSQAPTRPAPTTLSTTTGPAPTTPV
ATTVPAPTTPTPTPDLPSSSNSSVLPTPPATEAPSSPPPEYVCNCSVVGSLNVNRCNQTTGQCECRPGYQGLHC
ETCKEGFYLNYSGLCQPCDCSPHGALSIPCNR
```

**Important features of the protein:****Signal peptide:**

amino acids 1-25

**N-glycosylation sites.**

amino acids 30-33, 172-175, 195-198, 208-211, 235-238

**EGF-like domain cysteine pattern signature.**

amino acids 214-226.



**FIGURE 285**

AACACAGACGTTCCCTCGCGGCCCTGGCACCTCTAACCCACAGACATGCTGCTGCTGCTGCTGCCCTGCTCTGGGGG  
AGGGAGAGGGCGGAAGGACAGACAAGTAACTGCTGACGATGCAGAGTTCGGTGACGGTGCAGGAAGGCTGTGT  
GTCCATGTGCCCTGCTCCTTCTCCTACCCCTCGCATGGCTGGATTTACCTTGGCCCAGTAGTTTCATGGCTACTGG  
TTCCGGGAAGGGGCCAATACAGACCAGGATGCTCCAGTGGCCACAAACACCCAGCTCGGGCAGTGTGGAGGAG  
ACTCGGAGCCGATTCCACCTCCTTGGGGACCCACATACCAAGAAATTGCACCTGAGCATCAGAGATGCCAGAAGA  
AGTGATGCGGGGAGATCACTTCTTTCTGATGGAGAAAGGAAGTATAAAATGGAATTATAAACATCACCGGCTCTCT  
GTGAATGTGACAGCCTTGACCCACAGGCCAACATCCTCATCCAGGCACCCTGGAGTCCGGCTGCCCCAGAAT  
CTGACCTGCTCTGTGCCCTGGGCCTGTGAGCAGGGGACACCCCTATGATCTCCTGGATAGGGACCTCCGTGTCC  
CCCCTGGACCCCTCCACCACCCGCTCCTCGGTGCTCACCCCTCATCCCACAGCCCCAGGACCATGGCACAGCCCT  
ACCTGTCAGGTGACCTTCCCTGGGGCCAGCGTGACCAGAACAGACCGTCCATCTCAACGTGTCCTACCGCCCT  
CAGAACTTGACCATGACTGTCTTCCAAGGAGCAGGCACAGTATCCACAGCTTTGGGAAATGGCTCATCTCTGTCA  
CTCCCAGAGGGCCAGTCTCTGCGCCTGGTCTGTGCAGTTGATGCAGTTGACAGCAATCCCCCTGCCAGGCTGAGC  
CTGAGCTGGAGAGGCCTGACCCTGTGCCCTCACAGCCCTCAAACCCGGGGGTGCTGGAGCTGCCTTGGGTGCAC  
CTGAGGGATGCAGCTGAATTACCTGCAGAGCTCAGAACCCTCTCGGCTCTCAGCAGGTCTACCTGAACGTCTCC  
CTGCGAGCAGCAAAGCCACATCAGGAGTGACTCAGGGGTGGTTCGGGGGAGCTGGAGCCACAGCCCTGGTCTTCTGT  
TCCTTGTCCGTCTATCTTCGTTGTAGTGAGGTCCTGCGGAGGAAGAAATCGGCAAGGCCACAGCGGGCTGGGAGAT  
ACGGGATAGAGGATGCAAACGCTGTGAGGGTTGAGCCTCTCAGGGGCCCTGACTGAACCTTGGGCAGAAGAC  
AGTCCCCCAGACCAGCCTCCCCAGCTTCTGCCCGCTCCTCAGTGGGGGAAGGAGAGCTCCAGTATGCATCCCTC  
AGCTTCCAGATGGTGAAGCCTTGGGACTCGCGGGGACAGGAGGCCACTGACACCGAGTACTCGGAGATCAAGATC  
CACAGATGAGAAACTGCAGAGACTCACCTGATTGAGGGATCACAGCCCCTCCAGGCAAGGGAGAAGTCAGAGGC  
TGATTCTTGTAGAATTAAACAGCCCTCAACGTGATGAGCTATGATAACACTATGAATTATGTGCAGAGTGAAAGC  
ACACAGGCTTTAGAGTCAAAGTATCTCAAACCTGAATCCACACTGTGCCCTCCCTTTATTTTTTAACTAAAG  
ACAGACAAATTCCTTA

287/615

**FIGURE 286**

MLLLLLPLLWGRERAEGQTSKLLTMQSSVTVQEGLCVHVPCSFSPSHGWIYPGPVVHGYWFEREGANTDQDAPVA  
TNNPARAVWEETRDRFHLGDPHTKNCTLSIRDARRSDAGRYFFRMEKGSIKWNYKHHRLSVNVTALHTRPNILI  
PGTLESGCPQNLTCSVPWACEQGTTPPMISWIGTSVSPLDPSTTRSSVLTLIPQPQDHGTSLTQCQVTFPGASVTTN  
KTVHLNVSYPQNLTMTVFQGDGTVSTVLGNGSSLSLPEGQSLRLVCAVDAVDSNPPARLSLSWRGLTLCPSQPS  
NPGVLELPWVHLRDAAEFTCRAQNPLGSQQVYLNVSLSQSKATSGVTQGVVGGAGATALVFLSFCVIFVVVRSCRK  
KSARPAAGVGDTGIEDANAVRGSASQGPLEPWAEDSPPDQPPPASARSSVGEDELQYASLSFQMKPWDSRGQE  
ATDTEYSEIKIHR

**Signal peptide:**  
amino acids 1-15

**Transmembrane domain:**  
amino acids 351-370

288/615

**FIGURE 287**

CGCGAGCTGAGAGGAGCAGGTAGAGGGGCAGAGGCGGGACTGTCTGCTGGGGGAGCCGCCAGGAGGCTCCTCAG  
GCCGACCCACAGACCTGGCTGGCCAGGATGAAGTATCTCCGGCACCCGGCGGCCCAATGCCACCCCTATTCTGGCC  
ATCGGCGCTTTACCCCTCCTCCTCTTCAGTCTGCTAGTGTACACCCACCTGCAAGGTCCAGGAGCAGCCACCG  
GCGATCCCCGAGGCCCTGGCCTGGCCCACTCCACCCACCCGCCAGCCCCGGCCCCGTGCCATGCCAACACCTCT  
ATGGTCACCCACCCGGACTTCGCCACGCAGCCGCAGCACGTTTCAAACTTCCTCCTGTACAGACACTGCCGCCAC  
TTTCCCCTGCTGCAGGACGTGCCCCCTCTAAGTGC GCGCAGCCGGTCTTCCTGCTGCTGGTGATCAAGTCTCTCC  
CCTAGCAACTATGTGCGCCGCGAGCTGCTGCGGCGCACGTGGGGCCGCGAGCGCAAGGTACGGGGTTTGCAGCTG  
CGCCTCCTCTTCCTGGTGGGCACAGCCTCCAACCCGCACGAGGCCCGCAAGGTCAACCGGTGCTGGAGCTGGAG  
GCACAGACTCACGGAGACATCCTGCAGTGGGACTTCCACGACTCCTTCTTCAACCTCACGCTCAAGCAGGTCTCTG  
TTCTTACAGTGGCAGGAGACAAGGTGCGCCAACGCCAGCTTCGTGCTCAACGGGGATGATGACGTCTTTGCACAC  
ACAGACAACATGGTCTTCTACCTGCAGGACCATGACCCTGGCCGCCACCTCTTCGTGGGGCAACTGATCCAAAC  
GTGGGCCCCATCCGGGCTTTTGGAGCAAGTACTATGTGCCAGAGGTGGTGAAGTACAGCGGTACCCACCC  
TATTGTGGGGGTGGTGGCTTCTTGCTGTCCCGCTTACGGCCGCTGCCCTGCGCCGTGCTGCCCATGTCTTGGAC  
ATCTTCCCCATTGATGATGTCTTCTGGGTATGTGTCTGGAGCTTGAGGGACTGAAGCCTGCCTCCCACAGCGGC  
ATCCGCACGTCTGGCGTGCGGGCTCCATCGCAACACCTGTCTCCTTTGACCCCTGCTTCTACCGAGACCTGCTG  
CTGGTGCACCGCTTCTTACCTTATGAGATGCTGCTCATGTGGGATGCGCTGAACAGCCCCAACCTCACCTGCGGC  
AATCAGACACAGATCTACTGAGTCAGCATCAGGGTCCCCAGCCTCTGGGCTCCTGTTTCCATAGGAAGGGGCGAC  
ACCTTCTCTCCAGGAAGCTGAGACCTTTGTGGTCTGAGCATAAGGGAGTGCCAGGGAAGGTTTGAGGTTTGATGA  
GTGAATATTCTGGCTGGCGAACTCCTACACATCCTTCAAAACCCACCTGGTACTGTTCCAGCATCTTCCCTGGAT  
GGCTGGAGGAACTCCAGAAAATATCCATCTTCTTTTGTGGCTGCTAATGGCAGAAGTGCCTGTGCTAGAGTTCC  
AACTGTGGATGCATCCGTCCCGTTTGAGTCAAAGTCTTACTTCCCTGCTCTCACCTACTCACAGACGGGATGCTA  
AGCAGTGCACCTGCAGTGGTTTAAATGGCAGATAAGCTCCGTCTGCAGTTCAGGCCAGCCAGAACTCCTGTGTC  
CACATAGAGCTGACGTGAGAAATATCTTTCAGCCAGGAGAGAGGGGTCTGATCTTAACCCCTTCTCTGGGTCTC  
AGACAACTCAGAAGGTTGGGGGGATACCAGAGAGGTGGTGGAAATAGGACCGCCCCCTCCTTACTTGTGGGATCAA  
ATGCTGTAATGGTGGAGGTGTGGGCAGAGGAGGGAGGCAAGTGTCTTTGAAAGTTGTGAGAGCTCAGAGTTTCTG  
GGGTCTCATTAGGAGCCCCCATCCCTGTGTTCCCCAAGAATTACAGAGAACAGCACTGGGGCTGGAATGATCTTT  
AATGGGCCCAAGGCCAACAGGCATATGCCTCACTACTGCCTGGAGAAGGGAGAGATTAGGTCTCCAGCAGCCT  
CCCTACCCAGTATGTTTTACAGATTACGGGGGACCGGGTGAGCCAGTGACCCCTGCAGCCCCCAGCTTCAGG  
CCTCAGTGTCTGCCAGTCAAGCTTCACAGGCATTGTGATGGGGCAGCCTTGGGGAATATAAAATTTTGTGAAGAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

289/615

**FIGURE 288**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA65413

&lt;subunit 1 of 1, 372 aa, 1 stop

&lt;MW: 42515, pI: 8.92, NX(S/T): 6

MKYLRHRRPNATLILAIGAFILLFSLLVSPPTCKVQEQPPAIPALAWPTPPTRPAPAPCHANTSMVTHPDFAT  
QPQHVQNFLLYRHCRRHFLLQDVPPSKCAQPVFLLLVIKSSPSNYVRRELLRRTWGREKVRGLQLRLLFLVGTA  
SNPHEARKVNRLLLELEAQTHGDILQWDFHDSFFNLTLKQVLFQWQETRCANASFVLNGDDDVFAHTDNMVFYLQ  
DHDPGRHLEFVGQLIQNVGPPIRAFWISKYYVPEVVTQNERYPYCGGGGFLLSRFTAAALRRAHVLDIFPIDDVFL  
GMCLELEGLKPASHSGIRTSGVRAPSQHLSSFDPCFYRDLVVHRLPYEMLLMWDALNQPNLTCGNQTQIY

**Important features:****Type II transmembrane domain:**

Amino acids 15-34

**N-glycosylation sites:**

Amino acids 10-14;64-68;184-188;202-206;362-366;367-371

**TonB-dependent receptor proteins signature 1:**

Amino acids 1-32

**N-myristoylation sites:**

Amino acids 308-314;316-322

290/615

**FIGURE 289A**

CGCGCTCCCCGCGCGCCTCCTCGGGCTCCACGCGTCTTGCCCCGAGAGGCAGCCTCCTCCAGGAGCGGGGCCCT  
GCACACC**ATG**CCCCCGGGTGGGCAGGGGTGCGCGCCGCCGTGCGCGCCCGCCTGGCGCTGGCCTTGGCGCTGGC  
GAGCGTCTGAGTGGGCCCTCCAGCCGTGCGCTGCCCCACCAAGTGTAACCTGCTCCGCTGCCAGCGTGGACTGCCA  
CGGGCTGGGCCCTCCGCGCGGTTCCTCGGGGCATCCCCGCAACGCTGAGCGCCTTGACCTGGACAGAAATAATAT  
CACCAGGATCACCAGATGGACTTCGCTGGGGCTCAAGAACCTCCGAGTCTTGCACTCTGGAAGACAACCAGGTCAG  
CGTCATCGAGAGAGGCGCCTTCCAGGACCTGAAGCAGCTAGAGCGACTGCGCCTGAACAAGAATAAGCTGCAAGT  
CCTTCCAGAATTGCTTTTCCAGAGCACGCCGAAGCTCACCAGACTAGATTTGAGTGAAAACCAGATCCAGGGGAT  
CCCGAGGAAGGCGTTCGCGCGCATCACCGATGTGAAGAACCTGCAACTGGACAACAACCACATCAGCTGCATTGA  
AGATGGAGCCTTCCGAGCGCTGCGCGATTTGGAGATCCTTACCCTCAACAACAACAACATCAGTCGCATCCTGGT  
CACCAGCTTCAACCACATGCCGAAGATCCGAACCTCTGCGCCTCCACTCCAACCACCTCTACTGCGACTGCCACCT  
GGCCTGGCTCTCGGATTGGCTGCGACAGCGACGGACAGTTGGCCAGTTCACACTCTGCATGGCTCCTGTGCATTT  
GAGGGGCTTCAACGTGGCGGATGTGCAAGAAGGAGTACGTGTGCCAGCCCCCACTCGGAGCCCCCATCCTG  
CAATGCCAACTCCATCTCCTGCCCTTCGCCCTGCACGTGCAGCAATAACATCGTGGACTGTGAGGAAAGGGCTT  
GATGGAGATTCTGCCAACTTGCCGGAGGGCATCGTCGAAATACGCCCTAGAACAGAACTCCATCAAAGCCATCCC  
TGCAGGAGCCTTACCCAGTACAAGAACTGAAGCGAATAGACATCAGCAAGAATCAGATATCGGATATTGCTCC  
AGATGCCTTCCAGGGCCTGAAATCACTCACATCGCTGGTCTGTATGGGAACAAGATCACCAGATTGCCAAGGG  
ACTGTTTGATGGGCTGGTGTCCCTACAGCTGCTCCTCCTCAATGCCAACAAGATCAACTGCCTGCGGGTGAACAC  
GTTTCAGGACCTGCAGAACCTCAACTTGCTCTCCTGTATGACAACAAGCTGCAGACCATCAGCAAGGGGCTCTT  
CGCCCCCTCTGCAGTCCATCCAGACACTCCACTTAGCCCCAAAACCCATTTGTGTGCGACTGCCACTTGAAGTGGCT  
GGCCGACTACCTCCAGGACAACCCCATCGAGACAAGCGGGGCCCGCTGCAGCAGCCCCGCGCCGACTCGCCAAACA  
GCGCATCAGCCAGATCAAGAGCAAGAAGTTCCGCTGCTCAGGCTCCGAGGATTACCGCAGCAGGTTTCCAGCAGCA  
GTGCTTCATGGACCTCGTGTGCCCGAGAAGTGTGCTGTGAGGGCAGGATTGTGGACTGCTCCAACCAGAAGCT  
GGTCCGATCCCAAGCCACCTCCCTGAATATGTACCGACCTGCGACTGAATGACAATGAGGTATCTGTTCTGGA  
GGCCACTGGCATCTTCAAGAAGTTGCCAACCTGCGGAAAAATAAATCTGAGTAACAATAAGATCAAGGAGGTGCG  
AGAGGGAGCTTTTCATGGAGCAGCCAGCGTGCAGGAGCTGATGCTGACAGGGAACCAGCTGGAGACCGTGACGG  
GCGCGTGTTCGCTGGCCTCAGTGGCCTCAAAACCTTGATGCTGAGGAGTAACTTGATCAGCTGTGTGAGTAATGA  
CACCTTTGCGGGCCTGAGTTCGGTGAGACTGCTGTCCCTCTATGACAATCGGATCACCACCATCACCCCTGGGGC  
CTTACCACGCTTGTCTCCCTGTCCACCATAAACCTCCTGTCCAACCCCTTCAACTGCAACTGCCACCTGGCCTG  
GCTCGGCAAGTGGTTGAGGAAGAGGCGGATCGTCAGTGGGAACCCCTAGGTGCCAGAAGCCATTTTCTCCTCAAGGA  
GATTTCCATCCAGGATGTGGCCATCCAGGACTTACCTGTGATGGCAACGAGGAGAGTAGCTGCCAGCTGAGCCC  
GCGCTGCCCCGAGCAGTGCACCTGTATGGAGACAGTGGTGCATGCAGCAACAAGGGGCTCCGCGCCCTCCCCAG  
AGGCATGCCCCAAGGATGTGACCGAGCTGTACCTGGAAGGAAACCACCTAACAGCCGTGCCAGAGAGCTGTCCGC  
CCTCCGACACCTGACGCTTATTGACCTGAGCAACAACAGCATCAGCATGCTGACCAATTACACCTTCACTAACAT  
GTCTACCTCTCCACTCTGATCCTGAGCTACAACCGGCTGAGGTGCATCCCCGTCCACGCCCTTCAACGGGCTGCG  
GTCCCTGCGAGTGCTAACCTCCATGGCAATGACATTTCCAGCGTTCTTGAAGGCTCCTTCAACGACCTCACATC  
TCTTTCCCATCTGGCGCTGGGAACCAACCCACTCCACTGTGACTGCAGTCTTCCGTGGCTGTGCGAGTGGGTGAA  
GGCGGGGTACAAGGAGCCTGGCATCGCCGCTGCAGTAGCCCTGAGCCCATGGCTGACAGGCTCCTGCTCACCAC  
CCCAACCCACCGCTTCCAGTGCAAAGGGCCAGTGGACATCAACATTGTGGCCAAATGCAATGCCTGCCTCTCCAG  
CCCGTGCAAGAATAACGGGACATGCACCCAGGACCTGTGGAGCTGTACCGCTGTGCTGCCCC

291/615

**FIGURE 289B**

TACAGCTACAAGGGCAAGGACTGCACTGTGCCCATCAACACCTGCATCCAGAACCCTGTGAGCATGGAGGCACC  
TGCCACCTGAGTGACAGCCACAAGGATGGGTTCAGCTGCTCCTGCCCTCTGGGCTTTGAGGGGCAGCGGTGTGAG  
ATCAACCCAGATGACTGTGAGGACAACGACTGCGAAAACAATGCCACCTGCGTGGACGGGATCAACAACCTACGTG  
TGTATCTGTCCGCCTAACTACACAGGTGAGCTATGCGACGAGGTGATTGACCACTGTGTGCCTGAGCTGAACCTC  
TGTCAGCATGAGGCCAAGTGCATCCCCCTGGACAAAGGATTGAGCTGCGAGTGTGTCCCTGGCTACAGCGGGAAG  
CTCTGTGAGACAGACAATGATGACTGTGTGGCCACAAGTGCCGCCACGGGGCCCAGTGGTGGACACAATCAAT  
GGCTACACATGCACCTGCCCCAGGGCTTCAGTGGACCTTCTGTGAACACCCCCACCCATGGTCTTACTGCAG  
ACCAGCCCATGCGACCAAGTACGAGTGCCAGAACGGGGCCCAGTGCATCGTGGTGCAGCAGGAGCCCACCTGCCGC  
TGCCCCACAGGCTTCGCCGGCCCCAGATGCGAGAAGCTCATCACTGTCAACTTCGTGGGCAAAGACTCCTACGTG  
GAACTGGCCTCCGCCAAGGTCCGACCCAGGCCAACATCTCCCTGCAGGTGGCCACTGACAAGGACAACGGCATC  
CTTCTCTACAAAGGAGACAATGACCCCCCTGGCACTGGAGCTGTACCAGGGCCACGTGCGGCTGGTCTATGACAGC  
CTGAGTTCCCTCCAACCACAGTGTACAGTGTGGAGACAGTGAATGATGGGCAGTTTACAGTGTGGAGCTGGTG  
ACGCTAAACCAGACCCTGAACCTAGTAGTGACAAAGGAACTCCAAAGAGCCTGGGGAAGCTCCAGAAGCAGCCA  
GCAGTGGGCATCAACAGCCCCCTCTACCTTGGAGGCATCCCCACCTCCACCGGCCTCTCCGCCTTGCGCCAGGGC  
ACGGACCGGCCTCTAGGCGGCTTCCACGGATGCATCCATGAGGTGCGCATCAACAACGAGCTGCAGGACTTCAAG  
GCCCTCCCACCACAGTCCCTGGGGGTGTCAACAGGCTGCAAGTCTGCACCGTGTGCAAGCACGGCCTGTGCCGC  
TCCGTGGAGAAGGACAGCGTGGTGTGCGAGTGCCGCCAGGCTGGACCGGCCCACTCTGCGACCAAGGAGGCCCG  
GACCCCTGCCTCGGCCACAGATGCCACCATGGAATGTGTGGCAACTGGGACCTCATACATGTGCAAGTGTGCC  
GAGGGCTATGGAGGGGACTTGTGTGACAACAAGAATGACTCTGCCAATGCCTGCTCAGCCTTCAAGTGTACCAT  
GGGCAGTGCCACATCTCAGACCAAGGGGAGCCCTACTGCCTGTGCCAGCCCGGCTTTAGCGGCGAGCACTGCCAA  
CAAGAGAATCCGTGCCTGGGACAAGTAGTCCGAGAGGTGATCCGCCGCCAGAAAGGTTATGCATCATGTGCCACA  
GCCTCCAAGGTGCCATCATGGAATGTGCTGGGGGCTGTGGGGCCCAGTGTGTCAGCCCACCCGCAGCAAGCGG  
CGGAAATACGTCTTCCAGTGCACGGACGGCTCCTCGTTTGTAGAAGAGGTGGAGAGACACTTAGAGTGCGGCTGC  
CTCGCGTGTTCCTAAGCCCCCTGCCCGCTGCCTGCCACCTCTCGGACTCCAGCTTGATGGAGTTGGGACAGCCAT  
GTGGGACCCCCTGGTGATTGAGCATGAAGGAAATGAAGCTGGAGAGGAAGGTAAAGAAGAAGAGAATATTAAAGTA  
TATTGTAAATAAACAAAAATAGAACTTAAAAAAAAAAAAAAAAAAAAAAAAA

292/615

**FIGURE 290**

MAPGWAGVGAAVRARLALALASVLSGPPAVACPTKCTCSAASVDCHGLGLRAVPRGIPRNAERLDDLDRNNITR  
ITKMDFAGLKNLRVLHLEDNQVSVIERGAFQDLKQLERLRNLNKNKLQVLPELLFQSTPKLTRLDLSENQIQGIPR  
KAFRGITDVKNLQLDNNHISCIEDGAFRALRDLEILTNNNNNISRLVTSFNHMPKIRTLRLHSNHLYCDCHLAW  
LSDWLRQRRTVGQFTLCMAPVHLRGFNVADVQKKEYVCPAPHSEPPSCNANSISCPSPCTCSNNIVDCRGKGLME  
IPANLPEGIVEIRLEQNSIKAIPAGFTQYKKLRIDISKNQISDIAPDAFQGLKSLTSLVLYGNKITEIAKGLF  
DGLVSLQLLLLNANKINCLRVNTFQDLQNLNLLSLYDNKLQTTISKGLFAPLQSIQTLHLAQNPFFVCDCHLKWLD  
YLQDNPIETSGARCSSPRRLANKRISQIKSKKFRCSGSEDIRSRFSSECFMDLVCPEKCRCEGTIVDCSNQKLVR  
IPSHLPEYVTDLRLNDNEVSVLEATGIFKKLPNLRKINLSNNKIKEVREGAFDGAASVQELMLTGNQLETVHGRV  
FRGLSGLKTLMLRSNLISCVSNDTFAGLSSVRLLSLYDNRIITITPGAFTTLVSLSTINLLSNPFNCNCHLAWLG  
KWLRKRRIVSGNPRCQKPFLLKEIPIQDVAIQDFTCDGNEESSCQLSPRCPEQCTCMETVVRC SNKGLRALPRGM  
PKDVTELYLEGNHILTAVPRELSALRHLLTIDLSNNSISMLTNYTFSNMSHLSTLILSYNRLRCIPVHAFNGLRSL  
RVLTLLHGNDISSVPEGSFNDLTSLSHLALGTNPLHCDCSLRWLSEWVKAGYKEPGIARCSSPEPMADRLLLLTPT  
HRFQCKGPVDINIVAKCNACLSSPCKNNGTCTQDPVELYRCACPYSYKGDCTVPINTCIQNPCQHGGETCHLSDS  
HKDGFSCSCPLGFEGQRCINPDDCEDNDCENNATCVDGINNYVCICPPNYTGELCDEVIDHCVPENLNCQHEAK  
CIPLDKGFSCCECVPGYSGKLCETDNDDCVAHKCRHGAQCVDTINGYTCTCPQGFGSPFCEHPPPMVLLQTSPCDQ  
YECQNGAQCIIVVQOEPTCRCPPGFAGPRCEKLITVNFVGKDSYVELASAKVRPQANISLQVATDKDNGILLYKGD  
NDPLALELYQGHVRLVYDSLSSPPTTVYSVETVNDGQFHSVELVTNLNQLTNLVVDKGTPKSLGKLQKQPAVGINS  
PLYLGGIPTSTGLSALRQGTDRPLGGFHGCIHEVRINNELQDFKALPPQSLGVSPGCKSCTVCKHGLCRSVEKDS  
VVCECRPGWTGPLCDQEARDPCLGHRCHHGKCVATGTSYMCKCAEGYGGDLCDNKND SANACSAFKCHHGQCHIS  
DQGEPLYCLCQPGFSGEHCQQENPCLGQVVREVIRROKGYASCATASKVPTIMECRGGCGPQCCQPTRSKRRKYVFQ  
CTDGSSFEVEVERHLECGCLACS

**Signal peptide:**

amino acids 1-27

293/615

**FIGURE 291**

GGATGCAGGACGCTCCCCTGAGCTGCCTGTCACCGACTAGGTGGAGCAGTGTTTCTTCCGCAGACTCAACTGAGA  
AGTCAGCCTCTGGGGCAGGCACCAGGAATCTGCCTTTTTCAGTTCTGTCTCCGGCAGGCTTTGAGGATGAAGGCTG  
CGGGCATTCTGACCCTCATTGGCTGCCTGGTCACAGGCGCCGAGTCCAAAATCTACACTCGTTGCAAACCTGGCAA  
AAATATTCTCGAGGGCTGGCCTGGACAATTACTGGGGCTTCAGCCTTGAAACTGGATCTGCATGGCATATTATG  
AGAGCGGCTACAACACCACAGCCCCGACGGTCCTGGATGACGGCAGCATCGACTATGGCATCTTCCAGATCAACA  
GCTTCGCGTGGTGACGACGCGGAAAGCTGAAGGAGAACAACCACTGCCATGTGCCTGCTCAGCCTTGATCACTG  
ATGACCTCACAGATGCAATTATCTGTGCCAGGAAAATTGTTAAAGAGACACAAGGAATGAATATTGGCAAGGCT  
GGAAGAAACATTGTGAGGGCAGAGACCTGTCCGAGTGGAAAAAGGCTGTGAGGTTTCCTAAACTGGAACCTGGAC  
CCAGGATGCTTTGCAGCAACGCCCTAGGATTTGCAGTGAATGTCCAAATGCCTGTGTCTATCTTGTCCCGTTTCCT  
CCCAATATTCCTTCTCAAACCTGGAGAGGGAAAATTAAGCTATACTTTTAAGAAAATAAATATTCCATTTAAATGTC



294/615

**FIGURE 292**

MKAAGILTLIGCLVTGAESKIYTRCKLAKIFSRAGLDNYWGFSLGNWICMAYYESGYNTTAPTVLDDGSIDYGIF  
QINSAWCRRGKLENNHCHVACSALITDDLTDALICARKIVKETQGMNYWQGWKKHCEGRDLSEWKKGCEVS

**Signal peptide:**  
amino acids 1-19

295/615

**FIGURE 293**

AGAAAGCTGCACTCTGTTGAGCTCCAGGGCGCAGTGGAGGGAGGGAGTGAAGGAGCTCTCTGTACCCAAGGAAAG  
TGCAGCTGAGACTCAGACAAGATTACA**ATGA**ACCAACTCAGCTTCCTGCTGTTTCTCATAGCGACCACCAGAGGA  
TGGAGTACAGATGAGGCTAATACTTACTTCAAGGAATGGACCTGTTCTTCGTCTCCATCTCTGCCCAGAAGCTGC  
AAGGAAATCAAAGACGAATGTCCTAGTGCATTTGATGGCCTGTATTTTCTCCGCACTGAGAATGGTGTTATCTAC  
CAGACCTTCTGTGACATGACCTCTGGGGGTGGCGGCTGGACCCCTGGTGGCCAGCGTGCATGAGAATGACATGCGT  
GGGAAGTGCACGGTGGGCGATCGCTGGTCCAGTCAGCAGGGCAGCAAAGCAGACTACCCAGAGGGGGACGGCAAC  
TGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCCTGGCTACTACGAC  
ATCCAGGCCAAGGACCTGGGCATCTGGCACGTGCCCCAATAAGTCCCCCATGCAGCACTGGAGAAACAGCTCCCTG  
CTGAGGTACCGCACGGACACTGGCTTCCTCCAGACACTGGGACATAATCTGTTTGGCATCTACCAGAAATATCCA  
GTGAAATATGGAGAAGGAAAGTGTGGACTGACAACGGCCCGGTGATCCCTGTGGTCTATGATTTTGGCGACGCC  
CAGAAAACAGCATCTTATTACTCACCCATGAGCCAGCGGGAATCACTGCGGGATTGTTCAGTTCAGGGTATTT  
AATAACGAGAGAGCAGCCAACGCCTTGTGTGCTGGAATGAGGGTCACCGGATGTAACACTGAGCATCACTGCATT  
GGTGGAGGAGGATACTTTCCAGAGGCCAGTCCCCAGCAGTGTGGAGATTTTCTGGTTTTGATTGGAGTGGATAT  
GGAACTCATGTTGGTTACAGCAGCAGCCGTGAGATAACTGAGGCAGCTGTGCTTCTATTCTATCGTT**TGAG**AGTTT  
TGTGGGAGGGAACCCAGACCTCTCCTCCCAACCATGAGATCCCAAGGATGGAGAACAACTTACCCAGTAGCTAGA  
ATGTTAATGGCAGAAGAGAAAACAATAAATCATATTGACTCAAGAAAAAAA

296/615

**FIGURE 294**

MNQLSFLFLIATTRGWSTDEANTYFKEWTCSSSPSLPRSCKEIKDECPSAFDGLYFLRTENGVIYQTFCDMTSG  
GGGWTLVASVHENDMRGKCTVGDRWSSQQGSKADYPEGDGNWANYNTFGSAEAATSDDYKNPGYYDIQAKDLGIW  
HVPNKSPMQHWRNSSLLRYRTDTGFLQTLGHNLFGIYQKYPVKYGEKGCWTDNGPVIIPVVYDFGDAQKTASYSP  
YQOREFTAGFVQFRVFNNERAANALCAGMRVTGCNTEHHCIGGGYFPEASFPQCGDFSGFDWSGYGTHVGYSS  
REITEAAVLLFYR

297/615

**FIGURE 295**

CAGGCCATTTGCATCCCCTGTCCTTGTGTTTCGGAGCCAGGCCACACCGTCCTCAGCAGTGTGCATGTGTTAAAAA  
CGCCAAGCTGAATATATC**ATG**CCCCCTATTAAAACTTGTACATGGCTCCCCATTGGTTTTTGGAGAAAAGTTC AAG  
CTTTTACCTTGGTGTCTGCCTGTATCCCAGTGTTCAGGCTGGCTAGACGGCGGAAGAAGATCCTATTTTACTGT  
CACTTCCCAGATCTGCTTCTCACCAGAGAGATTCTTTTCTTAAACGACTATACAGGGCCCCAATTGACTGGATA  
GAGGAATACACCACAGGCATGGCAGACTGCATCTTAGTCAACAGCCAGTTCACAGCTGCTGTTTTTAAAGGAAACA  
TTCAAGTCCCTGTCTCACATAGACCCTGATGTCTCTATCCATCTCTAAATGTCACCAGCTTTGACTCAGTTGTT  
CCTGAAAAGCTGGATGACCTAGTCCCCAAGGGGAAAAAATTCTGTGCTCTCCATCAACAGATACGAAAGGAAG  
AAAAATCTGACTTTGGCACTGGAAGCCCTAGTACAGCTGCGTGGAAGATTGACATCCCAAGATTGGGAGAGGGTT  
CATCTGATCGTGGCAGGTGGTTATGACGAGAGAGTCTTGGAGAATGTGGAACATTATCAGGAATTGAAGAAAATG  
GTCCAACAGTCCGACCTTGGCCAGTATGTGACCTTCTTGAGGTCTTTCTCAGACAAACAGAAAATCTCCCTCCTC  
CACAGCTGCACGTGTGTGCTTTACACACCAAGCAATGAGCACTTTGGCATTGTCCCTCTGGAAGCCATGTACATG  
CAGTGCCCAAGTCATTGCTGTTAATTCGGGTGGACCCTTGGAGTCCATTGACCACAGTGTACAGGGTTTCTGTGT  
GAGCCTGACCCGGTGCATTCTCAGAAGCAATAGAAAAGTTTCATCCGTGAACCTTCTTTAAAAGCCACCATGGGC  
CTGGCTGGAAGAGCCAGAGTGAAGGAAAAATTTTCCCTGAAGCATTTACAGAACAGCTCTACCGATATGTTACC  
AAACTGCTGGTAT**TAA**TCAGATTGTTTTTAAAGATCTCCATTAAATGTCATTTTTATGGATTGTAGACCCAGTTTTGA  
AACCAAAAAAGAAACCTAGAATCTAATGCAGAAGAGATCTTTAAAAAATAAACTTGAGTCTTGAATGTGAGCCA  
CTTTCTATATACACACCTCCCTGTCCACTTTTCAGAAAAACCATGTCTTTTATGCTATAATCATTCCAAATTT  
TGCCAGTGTAAAGTTACAAATGTGGTGTCAATCCATGTTTACAGAGATATTTTAATTATATTTTCTCGGGATTAT  
TGCTCTTCTGTCTATAAATTTTGAATGATACTGTGCCTTAATTGGTTTTTCATAGTTTAAAGTGTGTATCATTATCA  
AAGTTGATTAATTTGGCTTCATAGTATAATGAGAGCAGGGCTATTGTAGTTCCAGATTCAATCCACCGAAGTGT  
TCACTGTCATCTGTTAGGGAATTTTGTGTTGCTCTTTGCCTGGATCCATAGCGAGAGTGCTCTGTATTTTT  
TTTAAGATAATTTGTATTTTGCACACTGAGATATAATAAAAGGTGTTTATCATAAAAAAAAAAAAAAAAAAAA

298/615

**FIGURE 296**

MPLLKLVHGSPLVFGEKFKLETLVSACIPVFR LARRRKKILFYCHFPDLLLLTKRDSFLKRLYRAPIDWIEEYTTG  
MADCILVNSQFTAAVFKETFKSLSHIDPDVLYPSLNVT SFDSVVPEKLDDLVPKGKKFLLLSINRYERKKNLTLA  
LEALVQLRGRLTSQDWERVHLIVAGGYDERVLENVEHYQELKKMVQQSDLGQYVTF LRSFSDKQKISLLHSCTCV  
LYTPSNEHFGIVPLEAMYMOCFVIAVN SGGPLESIDHSVTGFLCEPDPVHFSEAIEKFIREPSLKATMGLAGRAR  
VKEKFSPEAFTEQLYRYVTKLLV

**Signal peptide:**

amino acids 1-15

299/615

**FIGURE 297**

GACTACGCCGATCCGAGACGTGGCTCCCTGGGCGGCAGAACCA**ATG**TTGGACTTCGCGATCTTCGCCGTTACCTTC  
TTGCTGGCGTTGGTGGGAGCCGTGCTCTACCTCTATCCGGCTTCCAGACAAGCTGCAGGAATTCAGGGATTACT  
CCAAC**TGA**AGAAAAAGATGGTAATCTTCCAGATATTGTGAATAGTGGAAAGTTGCATGAGTTCCTGGTTAATTTG  
CATGAGAGATATGGGCCTGTGGTCTCCTTCTGGTTTGGCAGGCGCCTCGTGGTTAGTTTGGGCACTGTTGATGTA  
CTGAAGCAGCATATCAATCCCAATAAGACATCGGACCCTTTTGAAACCATGCTGAAGTCATTATTAAGGTATCAA  
TCTGGTGGTGGCAGTGTGAGTGAAAACCATGAGGAAAAAATTGTATGAAAATGGTGTGACTGATTCTCTGAAG  
AGTAACTTTGCCCTCCTCCTAAAGCTTTCAGAAGAATTATTAGATAAATGGCTCTCCTACCCAGAGACCCAGCAC  
GTGCCCCCTCAGCCAGCATATGCTTGGTTTGTCTATGAAGTCTGTTACACAGATGGTAATGGGTAGTACATTTGAA  
GATGATCAGGAAGTCATTGCTTCCAGAAGAATCATGGCACAGTTTGGTCTGAGATTGGAAAAGGCTTTCTAGAT  
GGGTCACTTGATAAAAAACATGACTCGGAAAAACAATATGAAGATGCCCTCATGCAACTGGAGTCTGTTTTAAGG  
AACATCATAAAAGAACGAAAAGGAAGGAACCTTCAGTCAACATATTTTCATTGACTCCTTAGTACAAGGGAACCTT  
AATGACCAACAGATCCTAGAAGACAGTATGATATTTTCTCTGGCCAGTTGCATAATAACTGCAAAATTGTGTACC  
TGGGCAATCTGTTTTTTAACCACCTCTGAAGAAGTTCAAAAAAATTATATGAAGAGATAAACCAAGTTTTTGGAA  
AATGGTCTGTACTCCAGAGAAAATTGAGCAGCTCAGATATTGTCAGCATGTGCTTTGTGAAACTGTTCGAACT  
GCCAACTGACTCCAGTTTCTGCCCAGCTTCAAGATATTGAAGGAAAAATTGACCGATTTATTATTCCTAGAGAG  
ACCCTCGTCCTTTATGCCCTTGGTGTGGTACTTCAGGATCCTAATACTTGGCCATCTCCACACAAGTTTGATCCA  
GATCGGTTTGATGATGAATTAGTAATGAAAACCTTTTCCTCACTTGGATTCTCAGGCACACAGGAGTGTCCAGAG  
TTGAGGTTTGCATATATGGTGACCACAGTACTTCTTAGTGTATTGGTGAAGAGACTGCACCTACTTTCTGTGGAG  
GGACAGGTTATTGAAACAAAGTATGAACTGGTAACATCATCAAGGGAAGAAGCTTGGATCACTGTCTCAAAGAGA  
TAT**TAA**AATTTTATACATTTAAAATCATTGTTAAATTGATTGAGGAAAACAACCATTTAAAAAAAATCTATGTTG  
AATCCTTTTATAAACAGTATCACTTTGTAATATAAACACCTATTTGTACTTAA

300/615

**FIGURE 298**

MLDFAIFAVTFLLALVGAVLYLYPASRQAAGIPGITPTEEKDGNLDPDIVNSGSLHEFLVNLHERYGPVVSFWFGR  
RLVVS LGTV DVLKQHINPNKTS DPFETMLKSLRLRYQSGGGSVSENHMRKKLYENGVTDSLKSNFALLLKLSEELL  
DKWLSYPETQHVPLSQHMLGFAMKSVTQMVMGSTFEDDQEVIRFQKNHGTWVSEIGKGFLDGSLDKNMTRKKQYE  
DALMQLESVLRNIIKERKGRNFSQHIFIDSLVQGNLNDQQILEDSMIFSLASCIITAKLCTWAICFLTTSSEEVQK  
KLYEEINQVFGNGPVTPEKIEQLRYCQHVLCE TVRTAKLTPVSAQLQDIEGKIDRFIIPRETLVLYALGVVLQDP  
NTWPSPHKFDPDRFDDELVMKTFSSLGFSGTQECPELRFAYMVTTVLLSVLVKRLHLLSVEGQVIETKYELVTSS  
REEAWITVSKRY

**Signal peptide:**

amino acids 1-18

**Transmembrane domain:**

amino acids 271-290

301/615

**FIGURE 299**

CTAGATTTGTCGGCTTGCGGGGAGACTTCAGGAGTCGCTGTCTCTGAACTTCCAGCCTCAGAGACCGCCGCCCTT  
GTCCCCGAGGGCCATGGGCCGGGTCTCAGGGCTTGTGCCCTCTCGCTTCCTGACGCTCCTGGCGCATCTGGTGGT  
CGTCATCACCTTATTCTGGTCCCGGGACAGCAACATACAGGCCTGCCTGCCTCTCACGTTACCCCCGAGGAGTA  
TGACAAGCAGGACATTCAGCTGGTGGCCGCGCTCTCTGTCACCCTGGGCCTCTTTGCAGTGGAGCTGGCCGGTTT  
CCTCTCAGGAGTCTCCATGTTCAACAGCACCCAGAGCCTCATCTCCATTGGGGCTCACTGTAGTGCATCCGTGGC  
CCTGTCTTCTTCATATTCGAGCGTTGGGAGTGCCTACGTATTGGTACATTTTTGTCTTCTGCAGTGCCCTTCC  
AGCTGTCACTGAAATGGCTTTATTTCGTCACCGTCTTTGGGCTGAAAAAGAAACCTTCTGATTACCTTCATGACG  
GGAACCTAAGGACGAAGCCTACAGGGGCAAGGGCCGCTTCGTATTCCTGGAAGAAGGAAGGCATAGGCTTCGGTT  
TTCCCCTCGGAACTGCTTCTGCTGGAGGATATGTGTGGAATAATTACGTCTTGAGTCTGGGATTATCCGCATT  
GTATTTAGTGCTTTGTAATAAAATATGTTTTGTAGTAACATTAAGACTTATATACAGTTTTAGGGGACAATTAA  
AAAAAAAAA



302/615

**FIGURE 300**

MGRVSGLVPSRFLTLLAHLVVVITLFWSRDSNIQACLPLTFTPEEYDKQDIQLVAALSVTLGLFAVELAGFLSGV  
SMFNSTQSLISIGAHCSASVALSFFIFERWECTTYWYIFVFCSALPAVTEMALEFVTVFGLKKKPF

**Transmembrane domain:**

amino acids 12-28 (type II), 51-66, 107-124

303/615

**FIGURE 301**

CTGGGACCCCGAAAAGAGAAGGGGAGAGCGAGGGGACGAGAGCGGAGGAGGAAGATGCAACTGACTCGCTGCTGC  
TTCGTGTTCTCTGGTGCAGGGTAGCCTCTATCTGGTCATCTGTGGCCAGGATGATGGTCCTCCCGGCTCAGAGGAC  
CCTGAGCGTGATGACCACGAGGGCCAGCCCCGGCCCCGGGTGCCCTCGGAAGCGGGGCCACATCTCACCTAAGTCC  
CGCCCCATGGCCAATTCCACTCTCCTAGGGCTGCTGGCCCCGCCTGGGGAGGCTTGGGGCATTCTTGGGCAGCCC  
CCCAACCGCCCGAACCACAGCCCCCACCCTCAGCCAAGGTGAAGAAAATCTTTGGCTGGGGCGACTTCTACTCC  
AACATCAAGACGGTGGCCCTGAACCTGCTCGTCACAGGGAAGATTGTGGACCATGGCAATGGGACCTTCAGCGTC  
CACTTCCAACACAATGCCACAGGCCAGGGAAACATCTCCATCAGCCTCGTCCCCCAGTAAAGCTGTAGAGTTC  
CACCAGGAACAGCAGATCTTCATCGAAGCCAAGGCCTCCAAAATCTTCAACTGCCGGATGGAGTGGGAGAAGGTA  
GAACGGGGCCGCCGGACCTCGCTTTGCACCCACGACCCAGCCAAGATCTGCTCCCGAGACCACGCTCAGAGCTCA  
GCCACCTGGAGCTGCTCCCAGCCCTTCAAAGTCGTCTGTGTCTACATCGCCTTCTACAGCACGGACTATCGGCTG  
GTCCAGAAGGTGTGCCAGATTACAACTACCATAGTGATACCCCTACTACCCATCTGGGTGACCCGGGGCAGGC  
CACAGAGGCCAGGCCAGGGCTGGAAGGACAGGCCTGCCATGCAGGAGACCATCTGGACACCGGGCAGGGAAGGG  
GTTGGGCCTCAGGCAGGGAGGGGGGTGGAGACGAGGAGATGCCAAGTGGGGCCAGGGCCAAGTCTCAAGTGGCAG  
AGAAAGGGTCCCAAGTGCTGGTCCCAACCTGAAGCTGTGGAGTGACTAGATCACAGGAGCACTGGAGGAGGAGTG  
GGCTCTCTGTGCAGCCTCACAGGGCTTTGCCACGGAGCCACAGAGAGATGCTGGGTCCCCGAGGCCTGTGGGCAG  
GCCGATCAGTGTTGGCCCCAGATCAAGTCATGGGAGGAAGCTAAGCCCTTGGTTCTTGCCATCCTGAGGAAAGATA  
GCAACAGGGAGGGGGAGATTTTCATCAGTGTTGGACAGCCTGTCAACTTAGGATGGATGGCTGAGAGGGCTTCCTAG  
GAGCCAGTCAGCAGGGTGGGTGGGGCCAGAGGAGCTCTCCAGCCCTGCCTAGTGGGCGCCCTGAGCCCCCTGTCT  
GTGTGCTGAGCATGGCATGAGGCTGAAGTGGCAACCCTGGGGTCTTTGATGTCTTGACAGATTGACCATCTGTCT  
CCAGCCAGGCCACCCCTTTCCAAAATTCCTCTTCTGCCAGTACTCCCCCTGTACCACCCATTGCTGATGGCACA  
CCCATCCTTAAGCTAAGACAGGACGATTGTGGTCCTCCACACTAAGGCCACAGCCCATCCGCGTGCTGTGTGTC  
CCTCTTCCACCCCAACCCCTGCTGGCTCCTCTGGGAGCATCCATGTCCCGGAGAGGGGTCCCTCAACAGTCAGCC  
TCACCTGTACAGCCGGGGTTCTCCCGGATCTGGATGGCGCCGCCCTCTCAGCAGCGGGCACGGGTGGGGCGGGC  
CGGGCCGCAGAGCATGTGCTGGATCTGTTCTGTGTGTCTGTCTGTGGGTGGGGGAGGGGAGGGAAGTCTTGTGA  
AACCGCTGATTGCTGACTTTTGTGTGAAGAATCGTGTCTTGGAGCAGGAAATAAAGCTTGCCCCGGGGCA

304/615

**FIGURE 302**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66521

&gt;&lt;subunit 1 of 1, 252 aa, 1 stop

&gt;&lt;MW: 28127, pI: 8.91, NX(S/T): 5

MLTRCCFVFLVQGSLLVICGQDDGPPGSEDPERDDHEGQPRPRVPRKRGHISPKSRPMANSTLLGLLAPPGEA  
WGILGQPPNRPNHSPPPSAKVKKIFGWGDFYSNIKTVALNLLVTGKIVDHGNGTFSVHFQHNATGQGNISISLVP  
PSKAVEFHQEQQIFIEAKASKIFNCRMWEKVERGRRTSLCTHDPAKICSRDHAQSSATWSCSQPFKVVCVYIAF  
YSTDYRLVQKVC PDYNYHSDTPYYPSG

**Important features of the protein:****Signal peptide:**

amino acids 1-14

**N-glycosylation sites.**

amino acids 62-65, 127-130, 137-140, 143-146

**2-oxo acid dehydrogenases acyltransferase**

amino acids 61-71

305/615

**FIGURE 303**

CGGTGGCCATGACTGCGGCCGTGTTCTTCGGCTGCGCCTTCATTGCCTTCGGGCCTGCGCTCGCCCTTTATGTCT  
TCACCATCGCCATCGAGCCGTTGCGTATCATCTTCCTCATCGCCGGAGCTTTCTTCTGGTTGGTGTCTCTACTGA  
TTTCGTCCCTTGTTTGGTTCATGGCAAGAGTCATTATTGACAACAAAGATGGACCAACACAGAAATATCTGCTGA  
TCTTTGGAGCGTTTGTCTCTGTCTATATCCAAGAAATGTTCCGATTTGCATATTATAAACTCTTAAAAAAGCCA  
GTGAAGGTTTGAAGAGTATAAACCAGGTGAGACAGCACCCCTCTATGCGACTGCTGGCCTATGTTTCTGGCTTGG  
GCTTTGGAATCATGAGTGGAGTATTTTCCTTTGTGAATACCCTATCTGACTCCTTGGGGCCAGGCACAGTGGGCA  
TTCATGGAGATTCTCCTCAATTCTTCCTTTATTAGCTTTTCATGACGCTGGTCATTATCTTGCTGCATGTATTCT  
GGGGCATTGTATTTTTTGTGGCTGTGAGAAGAAAAAGTGGGGCATCCTCCTTATCGTTCTCCTGACCCACCTGC  
TGGTGTGAGCCAGACCTTCATAAGTTCTTATTATGGAATAAACCTGGCGTCAGCATTATAATCCTGGTGCTCA  
TGGGCACCTGGGCATTCTTAGCTGCGGGAGGCAGCTGCCGAAGCCTGAAACTCTGCCTGCTCTGCCAAGACAAGA  
ACTTTCTTCTTTACAACCAGCGCTCCAGATTAACCTCAGGGAACCAGCACTTCCCAAACCGCAGACTACATCTTTA  
GAGGAAGCACAACCTGTGCCTTTTTCTGAAAATCCCTTTTTCTGGTGGAATTGAGAAAGAAATAAACTATGCAGATA

306/615

**FIGURE 304**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66658
><subunit 1 of 1, 257 aa, 1 stop
><MW: 28472, pI: 9.33, NX(S/T): 0
MTAAVFFGCAFIAGFPALALYVFTIAIEPLRIIFLIAGAFFWLVSLLISSLVWFMARVIIDNKGPTQKYLLIFG
AFVSVYIQEMFRFAYYKLLKKASEGLKSINPGETAPSMRLAYVSGLGFGIMSGVFSFVNTLSDSLGP GTVGIHG
DSPQFFLYSAFMTLVIIILLHVFWGIVFFDGCEKKKWGILLIVLLTHLLVSAQTFISSYYGINLASAFIILVLMGT
WAFLAAGGSCRSLLKCLLCQDKNFLLYNQSR
```

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domains:**

amino acids 32-51, 119-138, 152-169, 216-235

**Glycosaminoglycan attachment site.**

amino acids 120-123

**Sodium:neurotransmitter symporter family protein**

amino acids 31-65

307/615

FIGURE 305

[illegible]

308/615

**FIGURE 306**

MRSTILLFCLLGSTRSLPQLKPALGLPPTKLAPDQGTLPNQQSNQVFPSLSLIPLTQM  
LTLGPDHLHLLNPAAGMTPGTQTHPLTLGGLNVQQQLHPHVLPIFVTQLGAQGTILSSEE  
LPQIFTSLIIHSLFPGGILPTSQAGANPDVQDGSLEPAGGAGVNPATQGTPAGRLETPSG  
TDDDFAVTTPAGIQRSTHAIEEATTESANGIQ

**Signal peptide:**

amino acids 1-16

309/615

**FIGURE 307**

CCGGGGGACATGAGGTGGATACTGTTTCATTGGGGCCCTTATTGGGTCCAGCATCTGTGGCCAAGAAAAATTTTTTG  
GGGACCAAGTTTTGAGGATTAATGTCAGAAATGGAGACGAGATCAGCAAATGAGTCAACTAGTGAATTCAAACA  
ACTTGAAGCTCAATTTCTGGAAATCTCCCTCCTCCTTCAATCGGCCTGTGGATGTCCTGGTCCCATCTGTCAATC  
TGCAGGCATTTAAATCCTTCCTGAGATCCCAGGGCTTAGAGTACGCAGTGACAATTGAGGACCTGCAGGCCCTTT  
TAGACAATGAAGATGATGAAATGCAACACAATGAAGGGCAAGAACGGAGCAGTAATAACTTCAACTACGGGGCTT  
ACCATTTCCCTGGAAGCTATTTACCACGAGATGGACAACATTGCCGCAGACTTTCCTGACCTGGCGAGGAGGGTGA  
AGATTGGACATTCGTTTGAAAACCGGCCGATGTATGTAAGTTTACGACTGGGAAAGGCGTGAGGCGGCCGG  
CCGTTTGGCTGAATGCAGGCATCCATTCCCAGAGTGGATCTCCAGGCCACTGCAATCTGGACGGCAAGGAAGA  
TTGTATCTGATTACCAGAGGGATCCAGCTATCACCTCCATCTTGGAGAAAATGGATATTTTCTGTTGCCTGTGG  
CCAATCCTGATGGATATGTGTATACTCAAACCTCAAACCGATTATGGAGGAAGACGCGGTCCCGAAATCCTGGAA  
GCTCCTGCAATTGGTGCTGACCCAAATAGAACTGGAACGCTAGTTTTGCAGGAAAGGGAGCCAGCGACAACCCCTT  
GCTCCGAAGTGATCCATGGACCCCAACGCAATTCCGAAGTGAGGTGAAATCAGTGGTAGATTTTATCCAAAAAC  
ATGGGAATTTCAAGGGCTTCATCGACCTGCACAGCTACTCGCAGCTGCTGATGTATCCATATGGGTACTCAGTCA  
AAAAGGCCCCAGATGCCGAGGAACTCGACAAGGTGGCGAGGCTTGCGGCCAAAGCTCTGGCTTCTGTGCGGGCA  
CTGAGTACCAAGTGGGTCCCACCTGCACCACTGTCTATCCAGCTAGCGGGAGCAGCATCGACTGGGCGTATGACA  
ACGGCATCAAATTTGCATTACATTTGAGTTGAGAGATACCGGGACCTATGGCTTCTCCTGCCAGCTAACCCAGA  
TCATCCCCACTGCAGAGGAGCGTGGCTGGGGCTGAAGACCATCATGGAGCATGTGCGGGACAACCTCTACTAGG  
CGATGGCTCTGCTCTGTCTACATTTATTTGTACCCACACGTGCACGCACTGAGGCCATTGTTAAAGGAGCTCTTT  
CCTACCTGTGTGAGTCAGAGCCCTCTGGGTTTGTGGAGCACACAGGCCCTGCCCTCTCCAGCCAGCTCCCTGGAG  
TCGTGTGTCCTGGCGGTGTCCTGCAAGAACTGGTTCTGCCAGCCTGCTCAATTTTGGTCTGCTGTTTTTGATG  
AGCCTTTTGTCTGTTTCTCCTTCCACCCTGCTGGCTGGGCGGCTGCACTCAGCATCACCCCTTCCCTGGGTGGCAT  
GTCTCTCTTACCTCATTTTTAGAACCAAGAACATCTGAGATGATTCTCTACCCTCATCCACATCTAGCCAAGC  
CAGTGACCTTGCTCTGGTGGCACTGTGGGAGACACCACTTGTCTTTAGGTGGGTCTCAAAGATGATGTAGAATTT  
CCTTTAATTTCTCGCAGTCTTCTGGAAAATATTTTCTTTGAGCAGCAAATCTTGTAGGGATATCAGTGAAGGT  
CTCTCCCTCCCTCCTCTCCTGTTTTTTTTTTTTTTTGGAGACAGAGTTTTGCTCTTGTGTTGCCAGGCTGGAGTGA  
TGGCTCGATCTTGGCTCACCACAACCTCTGCCTCCTGGGTTCAAGCAATTTCTCCTGCCTCAGCCTCTTGTAGT  
TTGGTTTATAGGCGCATGCCACCATGCCTGGCTAATTTTGTGTTTTTAGTAGAGACAGGGTTTCTCCATGTTGGT  
CAGGCTGGTCTCAAACCTCCCAACCTCAGGTGATCTGCCCTCCTTGGCCTCCAGAGTGCTGGGATTACAGGTGTG  
AGCCACTGTGCGGGGCGGTCCTCCTTTTTTAGGCCTGAATACAAAGTAGAAGATCACTTCTCTTCACTGTGC  
TGAGAATTTCTAGATACTACAGTTCTTACTCCTCTCTTCCCTTTGTTATTCACTGTGACCAGGATGGCGGGAGGG  
GATCTGTGTCACTGTAGGTACTGTGCCAGGAAGGCTGGGTGAAGTGACCATCTAAATTGCAGGATGGTGAATTT  
ATCCCCATCTGTCTAATGGGCTTACCTCCTCTTTGCCCTTTGAACTCACTTCAAAGATCTAGGCCCTCATCTTAC  
AGGTCCCTAAATCACTCATCTGGCCTGGATAATCTCACTGCCCTGGCACATTCCCATTTGTGCTGTGGTGTATCCT  
GTGTTTCTTGTCTGTTTGT  
TTTTGTATCCTGGACCACAAGTTCCCTAAGTAGAGCAAGAATTCATCAACCAGCTGCCTCTTGTTCATTTACCT  
CAGCACGTACCATCTGTCTTTTGT  
TCTTAACCTCCTGCCTAGGATTTGTACAGCATCTGGTGTGTGCTTATAAGCCAATAAATATTCAATGTGAAAAA  
AAAAAAAAAA



310/615

**FIGURE 308**

MRWILFIGALIGSSICGQEKFFGDQVLRINVRNGDEISKLSQLVNSNNLKLNFWKSPSSFNRPVDVLVPSVSLQA  
FKSFLRSQGLEAYVTIEDLQALLDNEDEMQHNEGQERSNNFNFGAYHSLEAIYHEMDNIAADFPDLARRVKIG  
HSFENRPMYVLKFSTGKGVRRPAVWLNAGIHSREWISQATAIWTARKIVSDYQRDPAITSILEKMDIFLLPVANP  
DGYVYTQTQNLWRKTRSRNPGSSCIGADPNRNWNASEFAGKGASDNPCSEVYHGPHANSEVEVKSVVDFIQKHGN  
FKGFIDLHSYSQLLMYPYGYSVKKAPDAEELDKVARLAAKALASVSGTEYQVGPTCTTVYPASGSSIDWAYDNGI  
KFAFTFELRDTGTYGFLLPANQIIPAEETWLGLKTIMEHVRDNL

**Signal peptide:**  
amino acids 1-16

311/615

**FIGURE 309**

GGCTGACCGTGCTACATTGCCTGGAGGAAGCCTAAGGAACCCAGGCATCCAGCTGCCCACGCCTGAGTCCAAGAT  
TCTTCCCAGGAACACAAACGTAGGAGACCCACGCTCCTGGAAGCACCAGCCTTTATCTCTTCACCTTCAAGTCCC  
CTTTCTCAAGAATCCTCTGTTCTTTGCCCTCTAAAGTCTTGGTACATCTAGGACCCAGGCATCTTGCTTTCCAGC  
CACAAAGAGACAGATGAAGATGCAGAAAGGAAATGTTCTCCTTATGTTTGGTCTACTATTGCATTTAGAAGCTGC  
AACAAATTCCAATGAGACTAGCACCTCTGCCAACACTGGATCCAGTGTGATCTCCAGTGGAGCCAGCACAGCCAC  
CAACTCTGGGTCCAGTGTGACCTCCAGTGGGGTCAGCACAGCCACCATCTCAGGGTCCAGCGTGACCTCCAATGG  
GGTCAGCATAGTCACCAACTCTGAGTTCATACAACCTCCAGTGGGATCAGCACAGCCACCAACTCTGAGTTCAG  
CACAGCGTCCAGTGGGATCAGCATAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCAC  
CAACTCTGAGTCCAGCACACCCTCCAGTGGGGCCAGCACAGTCACCAACTCTGGGTCCAGTGTGACCTCCAGTGG  
AGCCAGCACTGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAGGGCCAGCACTGCCACCAACTCTGAGTCTAG  
CACACTCTCCAGTGGGGCCAGCACAGCCACCAACTCTGACTCCAGCACAACTCCAGTGGGGCTAGCACAGCCAC  
CAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTAG  
GGCCAGCACTGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCCAGCACAGCCACCAACTCTGAGTCCAG  
AACGACCTCCAATGGGGCTGGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCCAGCACAGCCAC  
CAACTCTGACTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGG  
GGCCAGCACAGCCACCAACTCTGAGTCCAGCACGACCTCCAGTGGGGCTAGCACAGCCACCAACTCTGACTCCAG  
CACAACTCCAGTGGGGCCGGCACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGGGATCAGCACAGTCCAC  
CAATTCTGAGTCCAGCACACCCTCCAGTGGGGCCAAACACAGCCACCAACTCTGAGTCCAGTACGACCTCCAGTGG  
GGCCAAACACAGCCACCAACTCTGAGTCCAGCACAGTGTCCAGTGGGGCCAGCACTGCCACCAACTCTGAGTCCAG  
CACAACTCCAGTGGGGTCAGCACAGCCACCAACTCTGAGTCCAGCACAACTCCAGTGGGGCTAGCACAGCCAC  
CAACTCTGACTCCAGCACAACTCCAGTGAGGCCAGCACAGCCACCAACTCTGAGTCTAGCACAGTGTCCAGTGG  
GATCAGCACAGTCACCAATTCTGAGTCCAGCACAACTCCAGTGGGGCCAAACACAGCCACCAACTCTGGGTCCAG  
GAGTGAGGCAAAGCCTGGTGGGTCCCTGGTGCCGTGGGAAATCTTCTCATCACCTGGTCTCGGTTGTGGCGGC  
CGTGGGGCTCTTTGCTGGGTCTTCTTCTGTGTGAGAAACAGCCTGTCCCTGAGAAACACCTTTAACACAGCTGT  
CTACCACCTCATGGCCTCAACCATGGCCTTGGTCCAGGCCCTGGAGGGAATCATGGAGCCCCCACAGGCCAG  
GTGGAGTCTTAAGTGGTTCTGGAGGAGACAGTATCATCGATAGCCATGGAGATGAGCGGGAGGAACAGCGGGCC  
**CTGAG**CAGCCCCGGAAGCAAGTGCCGCATTCTTCAGGAAGGAAGAGACCTGGGCACCCAAGACCTGGTTTCCTTT  
CATTCATCCCAGGAGACCCCTCCAGCTTTGTTTGAGATCCTGAAAATCTTGAAGAAGGTATTCTCACCTTTCT  
TGCCTTTTACCAGACACTGGAAAGAGAATACTATATTGCTCATTTAGCTAAGAAATAAATACATCTCATCTAACAC  
ACACGACAAAGAGAAGCTGTGCTTGCCCCGGGGTGGGTATCTAGCTCTGAGATGAACTCAGTTATAGGAGAAAAC  
CTCCATGCTGGACTCCATCTGGCATTCAAATCTCCACAGTAAATCCAAAGACCTCAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

312/615

**FIGURE 310**

MKMQKGNVLLMFGLLLHLEAATNSNETSTSA NTGSSVSSGASTATNSGSSVTSSGVSTATIS  
GSSVTSNGVSI VTNSEFHTTSSGISTATNSEFSTASSGISIATNSESSTTSSGASTATNSESS  
TPSSGASTVTNSGSSVTSSGASTATNSESSTVSSRASTATNSESSTLSSGASTATNSDSSTTS  
SGASTATNSESSTTSSGASTATNSESSTVSSRASTATNSESSTTSSGASTATNSESRTTSNGA  
GTATNSESSTTSSGASTATNSDSSTVSSGASTATNSESSTTSSGASTATNSESSTTSSGASTA  
TNSDSSTTSSGAGTATNSESSTVSSGISTVTNSESSTPSSGANTATNSESSTTSSGANTATNS  
ESSTVSSGASTATNSESSTTSSGVSTATNSESSTTSSGASTATNSDSSTTSSEASTATNSESS  
TVSSGISTVTNSESSTTSSGANTATNSGSSVTSAGSGTAALTGMHTTSHSASTAVSEAKPGGS  
LVPWEIFLITLVSVVAAGLFA GLFFCVRNSLSLRNTFNTAVYHPHGLNHGLGPGPGGNHGAP  
HRPRWSPNWFWRFPVSSIAMEMSGRNSGP

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 510-532

313/615

## FIGURE 311A

CTAAGCCGGAGGATGTGCAGCTGCGGCGGGCGGCCGGCTACGAAGAGGACGGGGACAGGCGCCGTGCGAACCGA  
GCCCAGCCAGCCGGAGGACGCGGGCAGGGCGGGACGGGAGCCCGGACTCGTCTGCCGCCCGCTCGTCGCCGTCTCG  
TGCCCGCCCCCGCTCCCCGCGCGGAGCGGGAGGACCGCCACCTCGCGCCCGAGCCCGCTAGCGCGCGC  
CGGGCATGGTCCCCTCTTAAAGGCGCAGGCCGCGGGCGGGCGGGGTGTGCGGAACAAAGCGCCGGCGCGGGG  
CCTGCGGGCGGGCTCGGGGGCCGCGATGGGCGCGGGCGGGCCCGCGGGCGGGCGGGCGCTGCCCGGGCGGGGCTCG  
CGGCGCTAGGGCGGGCTGGCCTCCGTGGGCGGGGGCAGCGGGCTGAGGGCGCGCGGAGCCTGCGGCGGGCGGGCGG  
GGCGGGCGGGCGGGCCCGGGCGGGCGGAGCGGCGGGGCAATGGCGCGCGCGGGCGGGCGGCGCTGGCTCAGCGTGC  
TGCTCGGGCTCGTCTGGGCTTCGTGCTGGCCTCGCGGCTCGTCTGCCCGGGGCTTCGAGCTGAAGCGAGCGG  
GCCCAGGGCGCGCGCCAGCCCCGAGGGCTGCCGGTCCGGGAGGCGGGCGGCTTCCCAGGCCGGCGGGGCGCGCG  
GCGATGCGCGCGGGGCGCAGCTCTGCCCGCCCGGCTCGGACCCAGATGGCGGGCGGGCGCGGACAGGAACCTTCTCT  
TCGTGGGAGTCATGACCGCCAGAAATACCTGCAGACTCGGGCCGTGGCCGCTACAGAATGGTCCAAGACAA  
TTCCTGGGAAAGTTCAGTTCTTCTCAAGTGAGGGTCTGACACATCTGTACCAATTCCAGTAGTCCACTACGGG  
GTGTGGACGACTCCTACCCGCCCCAGAAGAAGTCTTCATGATGCTCAAGTACATGCACGACCACTACTTGGACA  
AGTATGAATGGTTTATGAGAGCAGATGATGACGTGTACATCAAAGGAGACCGTCTGGAGAACTTCTGAGGAGTT  
TGAACAGCAGCGAGCCCTCTTCTTGGGCAGACAGGCTGGGCACCACGGAAGAAATGGGAAACTGGCCCTGG  
AGCCTGGTGAGAACTTCTGCATGGGGGGGCTGGCGTGATCATGAGCCGGGAGGTGCTTCGGAGAAATGGTGCCGC  
ACATTGGCAAGTGTCTCCGGGAGATGTACACCACCCATGAGGACGTGGAGGTGGGAAGGTGTGTCCGGAGGTTTG  
CAGGGGTGCAGTGTGTCTGGTCTTATGAGATGCGGCAGCTTTTTTATGAGAATTACGAGCAGAACAAAAGGGGT  
ACATTAGAGATCTCCATAACAGTAAATTCACCAAGCTATCACATTACACCCCAACAAAACCCACCTACCAGT  
ACAGGCTCCACAGCTACATGCTGAGCCGCAAGATATCCGAGCTCCGCCATCGCACAAATACAGCTGCACCGCGAAA  
TTGTCTGATGAGCAATACAGCAACACAGAAATTCATAAAGAGGACCTCCAGCTGGGAATCCCTCCCTCTCTCA  
TGAGGTTTTCAGCCCCGCCAGCGAGAGGAGATTCTGGAATGGGAGTTTCTGACTGGAAATACCTTGTATTTCGGCAG  
TTGACGGCCAGCCCCCTCGAAGAGGAATGGACTCCGCCAGAGGGAAGCCTTGACGACATTGTATGCAGGTCA  
TGGAGATGATCAATGCCAACGCCAAGACCAGAGGGCGCATCATTGACTTCAAAGAGATCCAGTACGGCTACCGCC  
GGGTGAACCCCATGTATGGGGCTGAGTACATCCTGGACCTGCTGCTTCTGTACAAAAGCACAAAGGGGAAGAAAA  
TGACGGTCCCTGTGAGGAGGCACGCGTATTTACAGCAGACTTTCAGCAAAATCCAGTTTGTGGAGCATGAGGAGC  
TGGATGCACAAGAGTTGGCCAAGAGAATCAATCAGGAATCTGGATCCTTGTCTTCTCTCAAACTCCCTGAAGA  
AGCTCGTCCCCTTTCAGCTCCCTGGGTGGAAGAGTGAACAAAGAAACCCAAAGATAAAAAGATAAACATACTGA  
TTCCTTGTCTGGGCGTTTCGACATGTTTGTGAGATTTATGGGAAACTTTGAGAAGACGTGTCTTATCCCCAATC  
AGAACGTCAAGCTCGTGGTCTGCTTTTCAATTCTGACTCCAACCCTGACAAGGCCAAACAAGTTGAACGTATGA  
GAGATTACCGCATTAAGTACCCTAAAGCCGACATGCAGATTTTGCTGTGTCTGGAGAGTTTTCAAGAGCCCTGG  
CCCTGGAAGTAGGATCCTCCAGTTTAAACAATGAATCTTTGCTCTTCTTCTGCGACGTGACCTCGTGTTTACTA  
CAGAAATCCTTCAGCGATGTCGAGCAATACAGTTCTGGGCCAACAAATATATTTTCAATCATCTTCAGCCAGT  
ATGACCCAAAGATTGTTTATAGTGGGAAAGTTCCAGTGACAACCATTTTGCCTTTACTCAGAAACTGGCTTCT  
GGAGAAACTATGGGTTTGGCATCACGTGTATTTATAAGGGAGATCTTGTCCGAGTGGGTGGCTTTGATGTTTCCA  
TCCAAGGCTGGGGGCTGGAGGATGTGGACCTTTTCAACAAGGTTGTCCAGGCAGGTTTGAAGACGTTTAGGAGCC  
AGGAAGTAGGAGTAGTCCACGTCCACCATCCTGTCTTTGTGATCCCAATCTTGACCCCAACAGTACAAAATGT  
GCTTGGGTCCAAAGCATCGACCTATGGGTCCACCCAGCAGCTGGCTGAGATGTGGCTGGAAAAAATGATCCAA  
GTTACAGTAAAAGCAGCAATAATAATGGCTCAGTGAGGACAGCCTAATGTCCAGCTTTGCTGGAAAAGACGTTTT  
TAATTATCTAATTTATTTTTCAAAAATTTTTGTATGATCAGTTTTTGAAGTCCGTATACAAGGATATATTTTAC  
AAGTGGTTTTCTTACATAGGACTCCTTTAAGATTGAGCTTTCTGAACAAGAAGGTGATCAGTGTGTGCTTTGAA  
CACATCTTCTTGTGAACATTATGTAGCAGACCTGCTTAACCTTTGACTTGAATGTACCTGATGAACAAAACCTT  
TTTAAAAAATGTTTTCTTTTGAGACCCTTTGCTCCAGTCCATGGCAGAAAACGTGAACATTCCTGCAAAGTAT  
TATTGTAACAAAACACTGTAACCTCTGTTAAATGTTCTGTTGTGATTGTTAAACATTCCACAGATTCTACCTTTTGT  
GTTTTGTTTTTTTTTTTACAATTGTTTTAAAGCCATTTTCATGTTCCAGTTGTAAGATAAGGAAATGTGATAATA  
GCTGTTTCATCATTGTCTTCAGGAGAGCTTTCAGAGTTGATCATTTCTCTCATGGTACTCTGCTCAGCATGGC  
CAGTAGGTTTTTTGTTTGTGTTTTGTTCTTTTTTGAGACGGAGTCTCACTCTGTTACCCAGGCTGGAATG  
CAGTGGCGCAATCTGGCTCACTTTAACCTCACTTCCCTGGTTCAAGCAATTCCTGCTTTGCTTCCCGAGT  
AGCTGGGATTACAGGCACACACCACCGCCAGNTAGTTTTTTTGTATTTTGTAGAGACGGGGTTTTACCAT  
GCAAGCCCAGCTGGCCACGTAGGTTTTTAAAGCAAGGGGCGTGAAGAAGGCACAGTGAGGTATGTGGCTGTTCTCG  
TGGTAGTTCATTTCGGCTAAATAGACCTGGCATTAAATTTCAAGAAGGATTGGCATTCTCTTCTTGACCCTT  
CTCTTTAAAGGGTAAATATTAATGTTTAGAATGACAAAGATGAATTATTACAATAAATCTGATGTACACAGACT  
GAAACATACACACATACCCCTAATCAAAACGTTGGGGAAAAATGTATTTGGTTTTGTTCTTCTCATCTGTCTG  
TGTTATGTGGGTGGAGATGGTTTTTCAATCTTTCATTACTGTTTTGTTTTATCCTTTGTATCTGAAATACCTTTAA

314/615

**FIGURE 311B**

TTTATTTAATATCTGTTGTTTCAGAGCTCTGCCATTTCTTGAGTACCTGTTAGTTAGTATTATTTATGTGTATCGG  
GAGTGTGTTTAGTCTGTTTTATTTGCAGTAAACCGATCTCCAAAGATTTCTTTTGGAAACGCTTTTCCCCTCC  
TTAATTTTATATTCTTACTGTTTTACTAAATATTAAGTGTTCCTTGACAATTTGGTGCTCATGTGTTTTGGG  
GACAAAAGTGAAATGAATCTGTCATTATACCAGAAAGTTAAATTCTCAGATCAAATGTGCCTTAATAAATTTGTT  
TTCATTTAGATTTCAAACAGTGATAGACTTGCCATTTTAATACACGTCATTGGAGGGCTGCGTATTTGTAAATAG  
CCTGATGCTCATTTGGAAAAATAAACAGTGAACAATATTTTCTATTGTACTTTTCGAACCATTTTGTCTCATT  
ATTCCTGTTTAGCTGAAGAATTGTATTACATTTGGAGAGTAAAAAACTTAAACACGAAAAA

315/615

**FIGURE 312**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68836

&gt;&lt;subunit 1 of 1, 802 aa, 1 stop

&gt;&lt;MW: 91812, pI: 9.52, NX(S/T): 3

MAARGRRAWLSVLLGLVLGFVLASRLVLPRASELKRAGPRRRASPEGCRSGQAAASQAGGARG  
DARGAQLWPPGSDPDGGPRDRNFLFVGVMTAQKYLQTRAVAAYRTWSKTIPGKVQFFSSEGSD  
TSVPIPVVPLRGVDDSYPPQKKSFMLKYMHDHYLDKYEFMRADDDVYIKGDRLENFLRSLN  
SSEPLFLGQTGLGTTEEMGKLALPGENFCMGGPGVIMSREVLRRMVPHIGHKCLREMYTTHE  
VEVGRCVRRFAGVQCVWSYEMRQLFYENYEQNKKGYIRDLHNSKIHQAITLHPNKNPPYQYRL  
HSYMLSRKISELRHRTIQLHREIVLMSKYSNTEIHKEDLQLGIPPSFMRFPQRQREEILEWEE  
LTGKYLYSAVDGQPPRRGMDSAQREALDDIVMQVMEMINANAKTRGRIIDFKEIQYGYRRVNP  
MYGAEYILDLLLLLYKKHKGKKMTVPVRRHAYLQOTFSKIQFVEHEELDAQELAKRINQESGSL  
SFLSNSLKKLVFPQLPGSKSEHKEPKDKKINILIPLSGRFDMFVRFMGNFEKTCLIPNQNVKL  
VVLLFNSDSNPDKAKQVELMRDYRIKYPKADMQILPVSGEFSRALALEVGSSQFNNESLLFFC  
DVDLVFTTEFLQRCRANTVLGQQIYFPIIFSQYDPKIVYSGKVPSDNHFAFTQKTGFWRNYGF  
GITCIYKGDIVRVGGFDVSIQGWGLEDDVLFNKVVQAGLKTFERSQEVGVVHVHHPVFCDPNLD  
PKQYKMCLGSKASTYGSTQQLAEMWLEKNDPSYSKSSNNNGSVRTA

**Signal peptide:**

amino acids 1-23

316/615

**FIGURE 313**

GGCCGGACGCCTCCGCGTTACGGGATGAATTAACGGCGGGTTCCGCACGGAGGTTGTGACCCC  
TACGGAGCCCCAGCTTGCCACGCACCCCACCTCGGCGTCGCGCGGCGTGCCCTGCTTGTGACA  
GGTGGGAGGCTGGAACATATCAGGCTGAAAAACAGAGTGGGTACTCTCTTCTGGGAAGCTGGCA  
ACAAATGGATGATGTGATATATGCATTCCAGGGGAAGGGAAATTGTGGTGCTTCTGAACCCAT  
GGTCAATTAACGAGGCAGTTTCTAGCTACTGCACGTACTTCATAAAGCAGGACTCTAAAAGCT  
TTGGAATCATGGTGTGATGAAAGGGATTACTTTATACTGACTCTGTTTTGGGGAAGCTTTT  
TTGGAAGCATTTTTCATGCTGAGTCCCTTTTACCTTTGATGTTTGTAACCCATCTTGGTATC  
GCTGGATCAACAACCGCCTTGTGGCAACATGGCTCACCTACCTGTGGCATTATTGGAGACCA  
TGTTTGGTGTAAGTGATTATACTGGGGATGCATTTGTTCTGGAGAAAGAAGTGTCTATTA  
TCATGAACCATCGGACAAGAATGGACTGGATGTTCTGTGGAATTGCCTGATGCGATATAGCT  
ACCTCAGATTGGAGAAAATTTGCCTCAAAGCGAGTCTCAAAGGTGTTCTGGATTGTGGTGGG  
CCATGCAGGCTGCTGCCTATATCTTCATTCATAGGAAATGGAAGGATGACAAGAGCCATTTTCG  
AAGACATGATTGATTACTTTTGTGATATTCACGAACCACTTCAACTCCTCATATTTCCAGAAG  
GGACTGATCTCACAGAAAACAGCAAGTCTCGAAGTAATGCATTTGCTGAAAAAATGGACTTC  
AGAAATATGAATATGTTTTACATCCAAGAACTACAGGCTTTACTTTTGTGGTAGACCGTCTAA  
GAGAAGGTAAGAACCTTGATGCTGTCCATGATATCACTGTGGCGTATCCTCACACATTCCTC  
AATCAGAGAAGCACCTCCTCCAAGGAGACTTTCCAGGGAAATCCACTTTACGTCCACCGGT  
ATCCAATAGACACCTCCCCACATCCAAGGAGGACCTTCAACTCTGGTGCCACAAACGGTGGG  
AAGAGAAAGAAGAGAGGCTGCGTTCCTTCTATCAAGGGGAGAAGAATTTTATTTTACCGGAC  
AGAGTGTCAATCCACCTTGCAAGTCTGAACTCAGGGTCCTTGTGGTCAAATTGCTCTCTATAC  
TGTATTGGACCCTGTTTCAGCCCTGCAATGTGCCTACTCATATATTTGTACAGTCTTGTTAAGT  
GGTATTTTATAATCACCATTGTAATCTTTGTGCTGCAAGAGAGAATATTTGGTGGACTGGAGA  
TCATAGAACTTGCATGTTACCGACTTTTACACAAACAGCCACATTTAAATTCAAAGAAAAATG  
AGTAAGATTATAAGGTTTGCCATGTGAAAACCTAGAGCATATTTTGGAAATGTTCTAAACCTT  
TCTAAGCTCAGATGCATTTTTCATGACTATGTGCAATATTTCTTACTGCCATCATTATTTGT  
TAAAGATATTTTGCACCTTAATTTTGTGGGAAAAATATTGCTACAATTTTTTTTAACTCTCTGAA  
TGTAATTTTCGATACTGTGTACATAGCAGGAGTGATCGGGGTGAAATAACTTGGGCCAGAATA  
TTATTAAACAATCATCAGGCTTTTAAA

317/615

**FIGURE 314**

MHSRGREIVVLLNPWSINEAVSSYCTYFIKQDSKSFSGIMVSWKGIYFILTLFWGSFFGSIFML  
SPFLPLMFVNPSWYRWINNRLVATWLTLPVALLETMFGVKVIITGDAFVPGERSVIIMNHRTR  
MDWMFLWNCLMRYSYLRLEKICLKASLKGVPFGFGWAMQAAAYIFIHRKWKDDKSHFEDMIDYF  
CDIHEPLQLLI FPEGTDLTENSKSRNAFAEKNGLQKYEYVLHPRTTGFTFVVDRLREGKNLD  
AVHDITVAYPHNIPQSEKHLLOGDFPREIHFHVHRYPIDTLPTSKEDLQLWCHKRWEEKEERL  
RSFYQGEKNFYFTGQSVIPPCKSELRVLVVKLLSILYWTLFSPAMCLLIYLYSLVKWYFIITI  
VIFVLQERIFGGLEIIELACYRLLHKQPHLNSKKNE

**Important features of the protein:**

**Signal peptide:**

amino acids 1-22

**Transmembrane domains:**

amino acids 44-63, 90-108, 354-377



318/615

**FIGURE 315**

CGGCTCGAGCGGCTCGAGTGAAGAGCCTCTCCACGGCTCCTGCGCCTGAGACAGCTGGCCTGA  
CCTCCAAATCATCCATCCACCCCTGCTGTCTATCTGTTTTCATAGTGTGAGATCAACCCACAGG  
AATATCC**ATGG**CCTTTTGTGCTCATTTTGGTTCTCAGTTTCTACGAGCTGGTGTGAGGACAGTG  
GCAAGTCACTGGACCGGGCAAGTTTGTCCAGGCCTTGGTGGGGGAGGACGCCGTGTTCTCCTG  
CTCCCTCTTTCTGAGACCAGTGCAGAGGCTATGGAAGTGCGGTTCTTCAGGAATCAGTTCCA  
TGCTGTGGTCCACCTCTACAGAGATGGGGAAGACTGGGAATCTAAGCAGATGCCACAGTATCG  
AGGGAGAACTGAGTTTGTGAAGGACTCCATTGCAGGGGGGGCGTGTCTCTAAGGCTAAAAAA  
CATCACTCCCTCGGACATCGGCCTGTATGGGTGCTGGTTCAGTTCCCAGATTTACGATGAGGA  
GGCCACCTGGGAGCTGCGGGTGGCAGCACTGGGCTCACTTCCTCTCATTTCATCGTGGGATA  
TGTTGACGGAGGTATCCAGTTACTCTGCCTGTCTCAGGCTGGTTCCCCCAGCCACAGCCAA  
GTGGAAGGTCCACAAGGACAGGATTTGTCTTCAGACTCCAGAGCAAATGCAGATGGGTACAG  
CCTGTATGATGTGGAGATCTCCATTATAGTCCAGGAAAATGCTGGGAGCATATTGTGTTCCAT  
CCACCTTGCTGAGCAGAGTCATGAGGTGGAATCCAAGGTATTGATAGGAGAGACGTTTTTCCA  
GCCCTCACCTTGGCGCCTGGCTTCTATTTTACTCGGGTTACTCTGTGGTGCCCTGTGTGGTGT  
TGTCATGGGGATGATAATTGTTTTCTTCAAATCCAAAGGGAAAATCCAGGCGGAAGTGGACTG  
GAGAAGAAAGCACGGACAGGCAGAATTGAGAGACGCCCGGAAACACGCAGTGGAGGTGACTCT  
GGATCCAGAGACGGCTCACCCGAAGCTCTGCGTTTCTGATCTGAAAAGTGAACCCATAGAAA  
AGCTCCCCAGGAGGTGCCTCACTCTGAGAAGAGATTTACAAGGAAGAGTGTGGTGGCTTCTCA  
GGGTTTTCCAAGCAGGGAGACATTACTGGGAGGTGGACGTGGGACAAAATGTAGGGTGGTATGT  
GGGAGTGTGTCGGGATGACGTAGACAGGGGGAAGAACAATGTGACTTTGTCTCCCAACAATGG  
GTATTGGGTCCTCAGACTGACAACAGAACATTTGTATTTACATTTCAATCCCCATTTTATCAG  
CCTCCCCCCCCAGCACCCCTCCTACACGAGTAGGGGTCTTCTGGACTATGAGGGTGGGACCAT  
CTCCTTCTTCAATACAAATGACCAGTCCCTTATTTATACCCTGCTGACATGTCAGTTTGAAGG  
CTTGTTGAGACCCTATATCCAGCATGCGATGTATGACGAGGAAAAGGGGACTCCCATATTTCAT  
ATGTCCAGTGTCTTGGGGAT**TGAG**ACAGAGAAGACCCTGCTTAAAGGGCCCCACACCACAGACC  
CAGACACAGCCAAGGGAGAGTGCTCCCGACAGGTGGCCCCAGCTTCCTCTCCGGAGCCTGCGC  
ACAGAGAGTCACGCCCCCACTCTCCTTTAGGGAGCTGAGGTTCTTCTGCCCTGAGCCCTGCA  
GCAGCGGCAGTCACAGCTTCCAGATGAGGGGGGATTGGCCTGACCCTGTGGGAGTCAGAAGCC  
ATGGCTGCCCTGAAGTGGGGACGGAATAGACTCACATTAGGTTTAGTTTGTGAAAACCTCCATC  
CAGCTAAGCGATCTTGAACAAGTCACAACCTCCCAGGCTCCTCATTGTGCTAGTCACGGACAGT  
GATTCCTGCCCTCACAGGTGAAGATTAAAGAGACAACGAATGTGAATCATGCTTGCAGGTTTGA  
GGGCACAGTGTTTGTCTAATGATGTGTTTTTATATTATACATTTTCCCACCATAAACTCTGTTT  
GCTTATTCCACATTAATTTACTTTTCTCTATACCAAATCACCCATGGAATAGTTATTGAACAC  
CTGCTTTGTGAGGCTCAAAGAATAAAGAGGAGGTAGGATTTTCACTGATTCTATAAGCCCAG  
CATTACCTGATACCAAAACCAGGCAAAGAAAACAGAAGAAGAGGAAGGAAAACCTACAGGTCCA  
TATCCCTCATTAACACAGACACAAAATTTCTAAATAAAATTTTAACAAATTAACCTAAACAAT  
ATATTTAAAGATGATATATACTACTCAGTGTGGTTTGTCCCAAAATGCAGAGTTGGTTTAA  
TATTTAAATATCAACCAGTGTAATTCAGCACATTAATAAAGTAAAAAAGAAAACCATAAAAAA  
AAAAA

319/615

**FIGURE 316**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68866

&gt;&lt;subunit 1 of 1, 466 aa, 1 stop

&gt;&lt;MW: 52279, pI: 6.16, NX(S/T): 2

MAFVLILVLSFYELVSGWQVTGPGKFVQALVGEDAVFSCSLFPETSAEAMEVRFFRNQFHAV  
VHLYRDGEDWESKQMPQYRGRTEFVKDSIAGGRVSLRLKNITPSDIGLYGCWFSSQIYDEEAT  
WELRVAALGSLPLISIVGYVDGGIQLLCLSSGWFPQPTAKWKGPQGQDLSSDSRANADGYSLY  
DVEISIIIVQENAGSILCSIHAEQSHEVESKVLIGETFFQPSPWRLASILLGLLCGALCGVVM  
GMIIIVFFKSKGKIQAELDWRKKGQAELRDARKHAVEVTLDPETAHPKLCVSDLKTVTHRKAP  
QEVPHSEKRFRTRKSVVASQGFQAGRHYWEVDVGQNVGWYVGVCRDDVDRGKNNVTLSPNNGYW  
VLRLTTEHLYFTFNPHFISLPPSTPPTRVGVFLDYEGGTISFFNTNDQSLIYTLLTCQFEGLL  
RPYIQHAMYDEEKGTPIFICPVSWG

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 131-150, 235-259

320/615

**FIGURE 317**

GCACCTGCGACCACCGTGAGCAGTC**ATG**GCGTACTCCACAGTGCAGAGAGTCGCTCTGGCTTC  
TGGGCTTGTCCTGGCTCTGTGCTGCTGCTGCCCAAGGCCTTCCTGTCCCGCGGGAAGCGGCA  
GGAGCCGCCGCGACACCTGAAGGAAAATTGGGCCGATTTCCACCTATGATGCATCATCACCA  
GGCACCCCTCAGATGGCCAGACTCCTGGGGCTCGTTTTCCAGAGGTCTCACCTTGCCGAGGCATT  
TGCAAAGGCCAAAGGATCAGGTGGAGGTGCTGGAGGAGGAGGTAGTGGAAGAGGTCTGATGGG  
GCAGATTATTCCAATCTACGGTTTTGGGATTTTTTTTATATATACTGTACATTCTATTTAAGGT  
AAGTAGAATCATCCTAATCATATTACATCAAT**TGA**AAATCTAATATGGCGATAAAAATCATTTGT  
CTACATTAAAACCTTCTTATAGTTCATAAAATTATTTCAAATCCATCATCTCTTTAAATCCTGC  
CTCCTCTTCATGAGGTACTTAGGATAGCCATTATTTTCAGTTTCACATAAGAATGTTTACTCAA  
TGTTTAAGTGTTTTGCCCCAAAATTCACAATAACAAGGCAGAACTAGGACTTGAACATGGAT  
CTTTTGGTTCTTAATCCAGTGAGTGATACAATTCAATGCACTCCCCTGCCA

321/615

**FIGURE 318**

MAYSTVQRVALASGLVLALSLLLPKAFLSRGKRQEPPPTPEGKLGRFPPMMHHHQAPSDGQTF  
GARFQRSHLAEAFKAKGSGGGAGGGGSGRGLMGQIIPYGFIFLYILYILFKVSRILIILHQ

322/615

**FIGURE 319**

CCTTCACAGGACTCTTCATTGCTGGTTGGCAATGATGTATCGGCCAGATGTGGTGAGGGCTAG  
GAAAAGAGTTTGTGGGAACCTGGGTTATCGGCCTCGTCATCTTCATATCCCTGATTGTCTT  
GGCAGTGTGCATTGGACTCACTGTTTATTATGTGAGATATAATCAAAGAAGACCTACAATTA  
CTATAGCACATTGTCATTTACAACCTGACAACTATATGCTGAGTTTGGCAGAGAGGCTTCTAA  
CAATTTTACAGAAATGAGCCAGAGACTTGAATCAATGGTGAAAAATGCATTTTATAAATCTCC  
ATTAAGGGAAGAATTTGTCAAGTCTCAGGTTATCAAGTTCAGTCAACAGAAGCATGGAGTGTT  
GGCTCATATGCTGTTGATTTGTAGATTTCACTCTACTGAGGATCCTGAACTGTAGATAAAAT  
TGTTCAACTTGTTTTACATGAAAAGCTGCAAGATGCTGTAGGACCCCTAAAGTAGATCCTCA  
CTCAGTTAAAATTAAAAAATCAACAAGACAGAAACAGACAGCTATCTAAACCATTGCTGCGG  
AACACGAAGAAGTAAACTCTAGGTCAGAGTCTCAGGATCGTTGGTGGGACAGAAGTAGAAGA  
GGGTGAATGGCCCTGGCAGGCTAGCCTGCAGTGGGATGGGAGTCATCGCTGTGGAGCAACCTT  
AATTAATGCCACATGGCTTGTGAGTGCTGCTCACTGTTTTACAACATATAAGAACCCTGCCAG  
ATGGACTGCTTCCTTTGGAGTAACAATAAAACCTTCGAAAATGAAACGGGGTCTCCGGAGAAT  
AATTGTCCATGAAAAATACAAACACCCATCACATGACTATGATATTTCTCTTGCAGAGCTTTC  
TAGCCCTGTTCCCTACACAAATGCAGTACATAGAGTTTGTCTCCCTGATGCATCCTATGAGTT  
TCAACCAGGTGATGTGATGTTTGTGACAGGATTTGGAGCACTGAAAAATGATGGTTACAGTCA  
AAATCATCTTCGACAAGCACAGGTGACTCTCATAGACGCTACAACCTGCAATGAACCTCAAGC  
TTACAATGACGCCATAACTCCTAGAATGTTATGTGCTGGCTCCTTAGAAGGAAAAACAGATGC  
ATGCCAGGGTGACTCTGGAGGACCACTGGTTAGTTCAGATGCTAGAGATATCTGGTACCTTGC  
TGGAATAGTGAGCTGGGGAGATGAATGTGCGAAACCCAACAAGCCTGGTGTTTATACTAGAGT  
TACGGCCTTGCGGGACTGGATTACTTCAAAAACCTGGTATCTAAGAGACAAAAGCCTCATGGAA  
CAGATAACATTTTTTTTTTGTTTTTTGGGTGTGGAGGCCATTTTTAGAGATACAGAATTGGAGA  
AGACTTGCAAAACAGCTAGATTTGACTGATCTCAATAAACTGTTTGCTTGATGCATGTATTTT  
CTTCCCAGCTCTGTTCCGCACGTAAGCATCCTGCTTCTGCCAGATCAACTCTGTCATCTGTGA  
GCAATAGTTGAACTTTATGTACATAGAGAAATAGATAATACAATATTACATTACAGCCTGTA  
TTCATTTGTTCTCTAGAAGTTTGTGTCAGAATTTTGAAGTTGTTGACATAAATTTGTAATGCATA  
TATACAATTTGAAGCACTCCTTTTCTTCAGTTCTCAGCTCCTCTCATTTTCAGCAAATATCCA  
TTTTCAAGGTGCAGAACAAGGAGTGAAAGAAAATATAAGAAGAAAAAATCCCCTACATTTTA  
TTGGCACAGAAAAGTATTAGGTGTTTTTCTTAGTGGAATATTAGAAATGATCATATTCATTAT  
GAAAGGTCAAGCAAAGACAGCAGAATACCAATCACTTCATCATTTAGGAAGTATGGGAACATA  
GTTAAGGAAGTCCAGAAAGAAGCCAAGATATATCCTTATTTTCATTTCCAACAACACTACTATG  
ATAAATGTGAAGAAGATTCTGTTTTTTGTGACCTATAATAATTATACAAACTTCATGCAATG  
TACTTGTTCTAAGCAAATTAAAGCAAATATTTATTTAACATTGTTACTGAGGATGTCAACATA  
TAACAATAAAATATAAATCACCCA

323/615

**FIGURE 320**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68871

&gt;&lt;subunit 1 of 1, 423 aa, 1 stop

&gt;&lt;MW: 47696, pI: 8.96, NX(S/T): 3

MMYRPDVVRARKRVCWEPWVIGLVIFISLIVLAVCIGLTVHYVRYNQKKTYNYYSTLSFTTDK  
LYAEFGREASNNFTEMSQRLESMVKNAFYKSPLREEFVKSQVIKFSQQKHGVLAHMLLICRFH  
STEDPETVDKIVQLVLHEKLQDAVGPPKVDPHSVKIKKINKTETDSYLNHCCGTRRSKTLGQS  
LRIVGGTEVEEGEWPWQASLQWDGSHRCGATLINATWLVSAAHCFTTYKNPARWTASFGVTIK  
PSKMKRGLRRIIVHEKYKHPSHDYDISLAELSSVPYTNVHRVCLPDASYEFQPGDVMFVTG  
FGALKNDGYSQNHLRQAQVTLIDATTCNEPQAYNDAITPRMLCAGSLEGKTDACQGDSGGPLV  
SSDARDIWYLAGIVSWGDECAKPNKPGVYTRVTALRDWITSKTGI

**Transmembrane domain:**

amino acids 21-40 (type II)

324/615

**FIGURE 321**

CCGGGCTCCTGGGTGAGGCCGGCAAGTTTGGAGCGTGGTCAGACAATAGGGGCGTGGCTACGG  
CTCGCGGAGCGCAACCAACGCTCTAGACCAGACCTGGGCTCGAGACCATAACTGTTTGGCTTT  
AACAGTACGTGGGCGGCCGGAATCCGGGAGTCCGGTGACCCGGGCTGTGGTCTAGCATAAAGG  
CGGAGCCCAGAAGAAGGGGCGGGGTATGGGAGAAGCCTCCCCACCTGCCCCCGCAAGGCGGCA  
TCTGCTGGTCCTGCTGCTGCTCCTCTCTACCCTGGTGATCCCCCTCCGCTGCAGCTCCTATCCA  
TGATGCTGACGCCCAAGAGAGCTCCTTGGGTCTCACAGGCCTCCAGAGCCTACTCCAAGGCTT  
CAGCCGACTTTTTCCTGAAAGGTAACCTGCTTCGGGGCATAGACAGCTTATTCTCTGCCCCCAT  
GGACTTCCGGGGCCTCCCTGGGAATAACCAAAAGAGGAGAACCAGGAGCACCAGCTGGGGAA  
CAACACCCCTCTCCAGCCACCTCCAGATCGACAAGATGACCGACAACAAGACAGGAGAGGTGCT  
GATCTCCGAGAATGTGGTGGCATCCATTCAACCAGCGGAGGGGAGCTTCGAGGGTGATTTGAA  
GGTACCCAGGATGGAGGAGAAGGAGGCCCTGGTACCCATCCAGAAGGCCACGGACAGCTTCCAC  
ACAGAACTCCATCCCCGGGTGGCCTTCTGGATCATTAAGCTGCCACGGCGGAGGTCCCACCAG  
GATGCCCTGGAGGGCGGCCACTGGCTCAGCGAGAAGCGACACCGCCTGCAGGCCATCCGGGAT  
GGACTCCGCAAGGGGACCCACAAGGACGTCCTAGAAGAGGGGACCGAGAGCTCCTCCCACTCC  
AGGCTGTCCCCCGAAAGACCCACTTACTGTACATCCTCAGGCCCTCTCGGCAGCTGTAGGGG  
TGGGGACCGGGGAGCACCTGCCTGTAGCCCCCATCAGACCCTGCCCCAAGCACCATATGGAAA  
TAAAGTTCTTTCTTACATCTAAAAA

325/615

**FIGURE 322**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68879
><subunit 1 of 1, 242 aa, 1 stop
><MW: 27007, pI: 8.68, NX(S/T): 2
MGEASPPAPARRHLLVLLLLLSTLVIPSAAAPIHDADAQESSLGLTGLQSLQGFSLFL
KGNLLRGIDSLFSAPMDFRGLPGNYHKEENQEHQLGNNTLSSHLQIDKMTDNKTGEVLIS
ENVVASIQPAEGSFEGDLKVPRMEEKEALVPIQKATDSFHTELHPRVAFWIILPRRRSH
QDALEGGHWLSEKRHRLQAIRDGLRKGTHKDVLEEGTESSSHSRLSPRKTHLLYILRPSR
QL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-30

**N-glycosylation sites:**

Amino acids 97-101;112-116

**N-myristoylation sites:**

Amino acids 80-86;132-138;203-209;216-222



326/615

**FIGURE 323**

AGAGAAAGAAGCGTCTCCAGCTGAAGCCAATGCAGCCCTCCGGCTCTCCGCGAAGAAGTTCCC  
TGCCCCGATGAGCCCCCGCCGTGCGTCCCCGACTATCCCCAGGCGGGCGTGGGGCACC GGCC  
CAGCGCCGACGATCGCTGCCGTTTTGCCCTTGGGAGTAGGATGTGGTGAAAGGATGGGGCTTC  
TCCCTTACGGGGCTCACAAATGCCAGAGAAGATTCGGTGAAGTGTCTGCGCTGCCTGCTCTAC  
GCCCTCAATCTGCTCTTTTGGTTAATGTCCATCAGTGTGTTGGCAGTTTCTGCTTGGATGAGG  
GACTACCTAAATAATGTTCTCACTTTAACTGCAGAAACGAGGGTAGAGGAAGCAGTCATTTTG  
ACTTACTTTTCTGTGGTTCATCCGGTCATGATTGCTGTTTGTGCTGTTTCTTATCATTGTGGGG  
ATGTTAGGATATTGTGGAACGGTGAAAAGAAATCTGTTGCTTCTTGCATGGTACTTTTGAAGT  
TTGCTTGTCATTTTCTGTGTAGAACTGGCTTGTGGCGTTTGGACATATGAACAGGAACCTTATG  
GTTCCAGTACAATGGTCAGATATGGTCACTTTGAAAGCCAGGATGACAAATTATGGATTACCT  
AGATATCGGTGGCTTACTCATGCTTGGAAATTTTTTTTCAGAGAGAGTTTAAGTGCTGTGGAGTA  
GTATATTTCACTGACTGGTTGGAAATGACAGAGATGGACTGGCCCCCAGATTCCTGCTGTGTT  
AGAGAATTCCCAGGATGTTCCAAACAGGCCACCAGGAAGATCTCAGTGACCTTTATCAAGAG  
GGTGTGGGAAGAAAATGTATTCCTTTTTTGAGAGGAACCAACAACCTGCAGGTGCTGAGGTTT  
CTGGGAATCTCCATTGGGGTGACACAAATCCTGGCCATGATTCTCACCATTACTCTGCTCTGG  
GCTCTGTATTATGATAGAAGGGAGCCTGGGACAGACCAATGATGTCCTTGAAGAATGACAAC  
TCTCAGCACCTGTCAATGTCCTCAGTAGAACTGTTGAAACCAAGCCTGTCAAGAATCTTTGAA  
CACACATCCATGGCAAACAGCTTTAATACACACTTTGAGATGGAGGAGTTATAAAAAGAAATG  
TCACAGAAGAAAACCACAACTTGTTTTATTGGACTTGTGAATTTTTGAGTACATACTATGTG  
TTTCAGAAATATGTAGAAATAAAAATGTTGCCATAAAATAACACCTAAGCATATACTATTCTA  
TGCTTTAAATGAGGATGGAAAAGTTTCATGTCATAAGTCACCACCTGGACAATAATTGATGC  
CCTTAAATGCTGAAGACAGATGTCATACCCACTGTGTAGCCTGTGTATGACTTTTACTGAAC  
ACAGTTATGTTTTGAGGCAGCATGGTTTTGATTAGCATTTCCGCATCCATGCAAACGAGTCACA  
TATGGTGGGACTGGAGCCATAGTAAAGGTTGATTTACTTCTACCACTAGTATATAAAGTACT  
AATTAAATGCTAACATAGGAAGTTAGAAAATACTAATAACTTTTATTACTCAGCGATCTATTC  
TTCTGATGCTAAATAAATTATATATCAGAAAACCTTCAATATTGGTGACTACCTAAATGTGAT  
TTTTGCTGGTTACTAAAATATTCTTACCCTTAAAAGAGCAAGCTAACACATTGTCTTAAGCT  
GATCAGGGATTTTTTGTATATAAGTCTGTGTTAAATCTGTATAATTCAGTCGATTTTCAGTTCT  
GATAATGTTAAGAATAACCATTATGAAAAGGAAAATTTGTCCTGTATAGCATCATTATTTTTA  
GCCTTTCTGTAAATAAAGCTTTACTATTCTGTCTGGGCTTATATTACACATATAACTGTTA  
TTTAAATACTTAACCCTAATTTTGAAAATTACCAGTGTGATACATAGGAATCATTATTCAGA  
ATGTAGTCTGGTCTTTAGGAAGTATTAATAAGAAAATTTGCACATAACTTAGTTGATTCAGAA  
AGGACTTGATGCTGTTTTTCTCCCAAATGAAGACTCTTTTGGACACTAAACACTTTTTAAAA  
AGCTTATCTTTGCCTTCTCCAAACAAGAAGCAATAGTCTCCAAGTCAATATAAATTCTACAGA  
AAATAGTGTCTTTTTCTCCAGAAAATGCTTGTGAGAATCATTAACATGTGACAATTTAG  
AGATTCTTTGTTTTATTTCACTGATTAATATACTGTGGCAAATTACACAGATTATTAATTTT  
TTTACAAGAGTATAGTATATTTATTTGAAATGGGAAAAGTGCATTTTACTGTATTTTGTGTAT  
TTTGTATTATTTCTCAGAATATGGAAAGAAAATTAATGTGTCAATAAATATTTTCTAGAGAG  
TAA

327/615

**FIGURE 324**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68880

&gt;&lt;subunit 1 of 1, 305 aa, 1 stop

&gt;&lt;MW: 35383, pI: 5.99, NX(S/T): 0

MAREDSVKCLRCLLYALNLLFWLMSISVLAVSAWMRDYLNNVLTTLTAETRVEEAVILTYFPVV  
HPVMIAVCCFLIIVGMLGYCGTVKRNLLLLAWYFGSLLVIFCVELACGVWTYEQELMVPVQWS  
DMVTLKARMTNYGLPRYRWLTHAWNFFQREFKCCGVVYFTDWLEMTENDWPPDSCCVREFPGC  
SKQAHQEDLSLDLYQEGCGKKMYSFLRGTKQLQVLRFLGISIGVTQILAMILTITLLWALYYDR  
REPGTDQMMSLKNDNSQHLSCPSVELLKPSLSRIFEHTSMANSFNTHFEMEEL

**Signal peptide:**

amino acids 1-33

**Transmembrane domains:**

amino acids 12-35, 57-86, 94-114, 226-248

328/615

**FIGURE 325**

AGCAGTGCATTGCTGGAGCGAGGAGAAGCTCACGAATCAGCTGCAGGTCTCTGTTTTGAAAAA  
GCAGAGATACAGAGGCAGAGGAAAAGGGTGGACTCCTATGTGACCTGTTCTTAGAGCAAGACA  
ATCACCATCTGAATTCCAGAAGCCCTGTTTCATGGTTGGGGATATTTTCTCGACTGCATGGAAT  
CAGAAAGAAGCAAAAGGATGGGAAATGCCTGCATTCCCCTGAAAAGAATTGCTTATTTCCCTAT  
GTCTCTTATCTGCGCTTTTGGCTGACTGAGGGGAAGAAACCAGCGAAGCCAAAATGCCCTGCCG  
TGTGTACTTGTACCAAAGATAATGCTTTATGTGAGAATGCCAGATCCATTCCACGCACCGTTC  
CTCCTGATGTTATCTCATTATCCTTTGTGAGATCTGGTTTTACTGAAATCTCAGAAGGGAGTT  
TTTTATTACGCCATCGCTGCAGCTCTTGTTATTCACATCGAACTCCTTTGATGTGATCAGTG  
ATGATGCTTTTATTGGTCTTCCACATCTAGAGTATTTATTCATAGAAAACAACAACATCAAGT  
CAATTTCAAGACATACTTTCCGGGGACTAAAGTCATTAATTCACCTTGAGCCTTGCAAACAACA  
ATCTCCAGACACTCCCAAAGATATTTTCAAAGGCCTGGATTCTTTAACAAATGTGGACCTGA  
GGGTAATTCATTTAATTGTGACTGTAACTGAAATGGCTAGTGGAATGGCTTGGCCACACCA  
ATGCAACTGTTGAAGACATCTACTGCGAAGGCCCCCAGAATACAAGAAGCGCAAAATCAATA  
GTCTCTCCTCGAAGGATTTGATTGCATCATTACAGAATTTGCAAAGTCTCAAGACCTGCCTT  
ATCAATCATTGTCCATAGACACTTTTTCTTATTTGAATGATGAGTATGTAGTCATCGCTCAGC  
CTTTTACTGGAAAATGCATTTTCCTTGAATGGGACCATGTGGAAAAGACCTTCCGGAATTATG  
ACAACATTACAGGCACATCCACTGTAGTATGCAAGCCTATAGTCATTGAACTCAGCTCTATG  
TTATTGTGGCCCAGCTGTTTGGTGGCTCTCACATCTATAAGCGAGACAGTTTTGCAAATAAAT  
TCATAAAAATCCAGGATATTGAAATTCTCAAAATCCGAAAACCCAATGACATTGAAACATTCA  
AGATTGAAAACAACTGGTACTTTGTTGTTGCTGACAGTTCAAAAGCTGGTTTTACTACCATTTAC  
AAATGGAACGGAAACGGATTCTACTCCCATCAATCCTTACACGCGTGGTACAGGGACACTGAT  
GTGGAATATCTAGAAATAGTCAGAACACCTCAGACACTCAGAACGCCTCATTTAATTCTGTCT  
AGTAGTTCCCAGCGTCCTGTAATTTATCAGTGGAACAAAGCAACACAATTATTCATAACCAA  
ACTGACATTCCTAACATGGAGGATGTGTACGCAGTGAAGCACTTCTCAGTGAAAGGGGACGTG  
TACATTTGCTTGACAAGATTCATTGGTGATTCCAAAGTCATGAAATGGGGAGGCTCCTCGTTC  
CAGGATATTCAGAGGATGCCATCGCGAGGATCCATGGTGTTCAGCCTCTTCAAATAAATAAT  
TACCAATATGCAATTCTTGGAAGTGATTACTCCTTTACTCAAGTGATAACTGGGATGCAGAG  
AAAGCCAAATTTGTGAAATTTAGGAATTAATGTTTCAGGCACCAAGATCATTACACATGTG  
TCCATTAATAAGCGTAATTTTCTTTTGGCTCCAGTTTTAAGGGAAATACACAGATTTACAAA  
CATGTCATAGTTGACTTAAGCGCATGAGACACCAAATTCGTGGCTGCCATCAGAAATTTTCT  
ACAGTACATGACCCGGATGAACTCAATGCATGATGACTCTTCTTATCACACTTGCAAATGAAT  
GCCTTTCAAACATTGAGACTGCTAGAACCAAGCACTACCAGTATCTCCATCCTTAAGTGTCCA  
GTCCAGTGATGTGGGAAGTTACCTTTTATAAGACAAAATTTAATTGTGTAAGTGTCTTTGCA  
GTGAAGATGTGTAAATAAGCGTTTAAATGGTATCTGTTACTCCAAAAAGAAATATTAATATGTA  
CTTTTCCATTTATTTATTCATGTGTACAGAAACAACTGCCAAATAAATGTTTACATTTTCTT  
TCATA

329/615

**FIGURE 326**

```
>>/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68882
><subunit 1 of 1, 557 aa, 1 stop
><MW: 63818, pI: 8.61, NX(S/T): 3
MESERSKRMGNACIPLKRIAYFLCLLSALLLTEGKKPAKPKCPAVCTCTKDNALCENARS
IPRTVPPDVISLSFVRSGFTEISEGSFLFTPSLQLLLFTSNSFDVISDDAFIGLPHLEYL
FIENNNIKSISRHTFRGLKSLIHLSLANNNLQTLPKDIFKGLDSLTVNVDLRGNSFNCDCK
LKWLVEWLGHNTATVEDIYCEGPPEYKKRKINSLSSKDFDCIITEFAKSQDLPLYQSLSID
TFSYLNDEYVVIAQPFTGKCIFLEWDHVEKTFERNYDNITGTSTVVCKPIVIETQLYVIVA
QLFGGSHIYKRDSFANKFIKIQDIEILKIRKPNDIETFKIENNWFVVDSSKAGFTTIY
KWNGNGFYSHQSLHAWYRDTDVEYLEIVRTPQTLRTPHLILSSSSQRPVIYQWNKATQLF
TNQTDIPNMEDVYAVKHFSVKGDVYICLTRFIGDSKVMKWGGSSFQDIQRMPSRGSMVFQ
PLQINNYQYAILGSDYSFTQVYNWDAEKAKFVKFQELNVQAPRSFTHVSINKRNFLFASS
FKGNTQIYKHVIVDLISA
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-34

**Transmembrane domain:**

Amino acids 281-306

**N-glycosylation sites:**

Amino acids 192-196;277-281;422-426

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 310-314

**Tyrosine kinase phosphorylation sites:**

Amino acids 228-235;378-385

**N-myristoylation sites:**

Amino acids 172-178;493-499

**Amidation site:**

Amino acids 33-37

330/615

**FIGURE 327**

CCAAGGCCAGAGCTGTGGACACCTTATCCCACTCATCCTCATCCTCTTCCTCTGATAAAGCCC  
CTACCAGTGCTGATAAAGTCTTTCTCGTGAGAGCCTAGAGGCCTTAAAAAAAAAAGTGCTTGA  
AAGAGAAGGGGACAAAGGAACACCAGTATTAAGAGGATTTTCCAGTGTTTCTGGCAGTTGGTC  
CAGAAGG**ATG**CCTCCATTCCCTGCTTCTCACCTGCCTCTTCATCACAGGCACCTCCGTGTCACC  
CGTGGCCCTAGATCCTTGTTCTGCTTACATCAGCCTGAATGAGCCCTGGAGGAACACTGACCA  
CCAGTTGGATGAGTCTCAAGGTCCTCCTCTATGTGACAACCATGTGAATGGGGAGTGGTACCA  
CTTCACGGGCATGGCGGGAGATGCCATGCCTACCTTCTGCATACCAGAAAACCACTGTGGAAC  
CCACGCACCTGTCTGGCTCAATGGCAGCCACCCCTAGAAAGGCGACGGCATTGTGCAACGCCA  
GGCTTGTGCCAGCTTCAATGGGAACCTGCTGTCTCTGGAACACCACGGTGGAAGTCAAGGCTTG  
CCCTGGAGGCTACTATGTGTATCGTCTGACCAAGCCCAGCGTCTGCTTCCACGTCTACTGTGG  
TCATTTTTATGACATCTGCGACGAGGACTGCCATGGCAGCTGCTCAGATAACCAGCGAGTGCAC  
ATGCGCTCCAGGAACTGTGCTAGGCCCTGACAGGCAGACATGCTTTGATGAAAATGAATGTGA  
GCAAAACAACGGTGGCTGCAGTGAGATCTGTGTGAACCTCAAAAACCTACCGCTGTGAGTG  
TGGGGTTGGCCGTGTGCTAAGAAGTGATGGCAAGACTTGTGAAGACGTTGAAGGATGCCACAA  
TAACAATGGTGGCTGCAGCCACTCTTGCCCTGGATCTGAGAAAGGCTACCAGTGTGAATGTCC  
CCGGGGCCTGGTGTCTGTCTGAGGATAACCACACTTGCCAAGTCCCTGTGTTGTGCAAATCAAA  
TGCCATTGAAGTGAACATCCCCAGGGAGCTGGTTGGTGGCCTGGAGCTCTTCCTGACCAACAC  
CTCCTGCCCAGGAGTGTCCAACGGCACCCATGTCAACATCCTCTTCTCTCAAGACATGTGG  
TACAGTGGTCGATGTGGTGAATGACAAGATTGTGGCCAGCAACCTCGTGACAGGTCTACCCAA  
GCAGACCCCGGGGAGCAGCGGGGACTTCATCATCCGAACCAGCAAGCTCGTGATCCCGGTGAC  
CTGCGAGTTTCCACGCCTGTACACCATTTCTGAAGGATACGTTCCCAACCTTCGAAACTCCCC  
ACTGGAAATCATGAGCCGAAATCATGGGATCTTCCCATTCACCTCTGGAGATCTTCAAGGACAA  
TGAGTTTGAAGAGCCTTACCGGGAAGCTCTGCCCACCCTCAAGCTTCGTGACTCCCTCTACTT  
TGGCATTGAGCCCGTGGTGCACGTGAGCGGCTTGAAAGCTTGGTGGAGAGCTGCTTTGCCAC  
CCCCACCTCCAAGATCGACGAGGTCTGAAATACTACCTCATCCGGGATGGCTGTGTTTCAGA  
TGACTCGGTAAAGCAGTACACATCCCGGGATCACCTAGCAAAGCACTTCCAGGTCCCTGTCTT  
CAAGTTTGTGGGCAAAGACCACAAGGAAGTGTCTGCACTGCCGGGTCTTGTCTGTGGAGT  
GTTGGACGAGCGTTCCCGCTGTGCCAGGGTTGCCACCGGCGAATGCGTCTGGGGCAGGAGG  
AGAGGACTCAGCCGGTCTACAGGGCCAGACGCTAACAGGCGGCCCGATCCGCATCGACTGGGA  
GGAC**TAG**TTTCGTAGCCATACCTCGAGTCCCTGCATTGGACGGCTCTGCTCTTTGGAGCTTCTC  
CCCCACCGCCCTCTAAGAACATCTGCCAACAGCTGGGTTTCAGACTTCACACTGTGAGTTCAG  
ACTCCCAGCACCAACTCACTCTGATTCTGGTCCATTAGTGGGCACAGGTACAGCACTGCTG  
AACAAATGTGGCCTGGGTGGGGTTTCATCTTTCTAGGGTTGAAAATAAACTGTCCACCCAGAA  
AGACACTCACCCCATTTCCCTCATTTCTTTCCTACACTTAAATACCTCGTGTATGGTGAATC  
AGACCACAAAATCAGAAGCTGGGTATAATATTTCAAGTTACAAACCCTAGAAAAATTAAACAG  
TTACTGAAATTATGACTTAAATACCCAATGACTCCTTAAATATGTAAATTATAGTTATACCTT  
GAAATTTCAATTCAAATGCAGACTAATTATAGGGAATTTGGAAGTGTATCAATAAAACAGTAT  
ATAATTTT

331/615

**FIGURE 328**

MPPFLLLTCLFITGTSVSPVALDPCSAYISLNEPWRNTDHQLDESQGPPLCDNHVNGEWYHFT  
GMAGDAMPTFCIPENHCGTHAPVWLNGSHPLEGDGIVQRQACASFNGNCCLWNTTVEVKACPG  
GYVYRLTKPSVCFHVYCGHFYDIDCEDCHGSCSDTSECTCAPGTVLGPDRQTCFDENECEQN  
NGGCSEICVNLKNSYRCECGVGRVLRSDGKTCEDEVEGCHNNNGGCSHSLGSEKGYQCECPRG  
LVLSEDNHTCQVPVLCKSNAIEVNI PRELVGGLELFLTNTSCRGVSNNGTHVNI LFSLKTCGTV  
VDVVNDKIVASNLVTGLPKQTPGSSGDFIIRTSKLLIPVTCEFPRLYTISEGYVPNLRNSPLE  
IMSRNHGIFPFTLEIFKDNEFEOPYREALPTLKL RDSLYFGIEPVVHVSGLES LVESCFATPT  
SKIDEVLKYYLIRDGCVSDDSVKQYTSRDHLAKHFQVPVFKFVGKDHKEVFLHCRVLVCGVLD  
ERSRCAQGCHRRMRRGAGGEDSAGLQGQTLTGGPIRIDWED

**Important features of the protein:****Signal peptide:**

amino acids 1-16

**N-glycosylation sites.**

amino acids 89-93, 116-120, 259-263, 291-295, 299-303

**Tyrosine kinase phosphorylation sites.**

amino acids 411-418, 443-451

**N-myristoylation sites.**amino acids 226-232, 233-239, 240-246, 252-258, 296-302, 300-306,  
522-528, 531-537**Aspartic acid and asparagine hydroxylation site.**

amino acids 197-209

**ZP domain proteins.**

amino acids 431-457

**Calcium-binding EGF-like proteins.**

amino acids 191-212, 232-253

332/615

**FIGURE 329**

GAGAGAGGCAGCAGCTTGCTCAGCGGACAAGGATGCTGGGCGTGAGGGACCAAGGCCTGCCCT  
GCACTCGGGCCTCCTCCAGCCAGTGCTGACCAGGGACTTCTGACCTGCTGGCCAGCCAGGACC  
TGTGTGGGGAGGCCCTCCTGCTGCCCTTGGGGTGACAATCTCAGCTCCAGGCTACAGGGAGACC  
GGGAGGATCACAGAGCCAGC**ATG**TTACAGGATCCTGACAGTGATCAACCTCTGAACAGCCTCG  
ATGTCAAACCCCTGCGCAAACCCCGTATCCCCATGGAGACCTTCAGAAAGGTGGGGATCCCCA  
TCATCATAGCACTACTGAGCCTGGCGAGTATCATCATTTGTGGTTGTCCTCATCAAGGTGATTC  
TGGATAAATACTACTTCCTCTGCGGGCAGCCTCTCCACTTCATCCCGAGGAAGCAGCTGTGTG  
ACGGAGAGCTGGACTGTCCCTTGGGGGAGGACGAGGAGCACTGTGTCAAGAGCTTCCCCGAAG  
GGCCTGCAGTGGCAGTCCGCCTCTCCAAGGACCGATCCACACTGCAGGTGCTGGACTCGGCCA  
CAGGGAACTGGTTCTCTGCCTGTTTCGACAACCTTCACAGAAGCTCTCGCTGAGACAGCCTGTA  
GGCAGATGGGCTACAGCAGAGCTGTGGAGATTGGCCCAGACCAGGATCTGGATGTTGTTGAAA  
TCACAGAAAAACAGCCAGGAGCTTCGCATGCGGAACTCAAGTGGGCCCTGTCTCTCAGGCTCCC  
TGGTCTCCCTGCACTGTCTTGCCCTGTGGGAAGAGCCTGAAGACCCCCCGTGTGGTGGGTGGGG  
AGGAGGCCTCTGTGGATTCTTGCCCTTGGCAGGTCAGCATCCAGTACGACAAACAGCACGTCT  
GTGGAGGGAGCATCCTGGACCCCCACTGGGTCTCACGGCAGCCCACTGCTTCAGGAAACATA  
CCGATGTGTTCAACTGGAAGGTGCGGGCAGGCTCAGACAACTGGGCAGCTTCCCATCCCTGG  
CTGTGGCCAAGATCATCATATTGAATTCAACCCCATGTACCCCAAAGACAATGACATCGCCC  
TCATGAAGCTGCAGTTCCCACTCACTTTCTCAGGCACAGTCAGGCCCATCTGTCTGCCCTTCT  
TTGATGAGGAGCTCACTCCAGCCACCCCACTCTGGATCATTGGATGGGGCTTTACGAAGCAGA  
ATGGAGGGAAGATGTCTGACATACTGCTGCAGGCGTCAGTCCAGGTCATTGACAGCACACGGT  
GCAATGCAGACGATGCGTACCAGGGGGGAAGTCACCGAGAAGATGATGTGTGCAGGCATCCCGG  
AAGGGGGTGTGGACACCTGCCAGGGTGACAGTGGTGGGGCCCCTGATGTACCAATCTGACCAGT  
GGCATGTGGTGGGCATCGTTAGCTGGGGCTATGGCTGCGGGGGCCCGAGCACCCCAGGAGTAT  
ACACCAAGGTCTCAGCCTATCTCAACTGGATCTACAATGTCTGGAAGGCTGAGCTG**TAAT**GCT  
GCTGCCCCCTTTGCAGTGCTGGGAGCCGCTTCCTTCCTGCCCTGCCACCTGGGGATCCCCCAA  
AGTCAGACACAGAGCAAGAGTCCCCTTGGGTACACCCCTCTGCCCACAGCCTCAGCATTTCTT  
GGAGCAGCAAAGGGCCTCAATTCTGTAAGAGACCCCTCGCAGCCCAGAGGCGCCCAGAGGAAG  
TCAGCAGCCCTAGCTCGGCCACACTTGGTGCTCCCAAGCATCCCAGGGAGAGACACAGCCCACT  
GAACAAGGTCTCAGGGGTATTGCTAAGCCAAGAAGGAACCTTCCCACACTACTGAATGGAAGC  
AGGCTGTCTTGTAAGGCCAGATCACTGTGGGCTGGAGAGGAGAAGGAAAGGGTCTGCGCCA  
GCCCTGTCCGTCTTCACCCATCCCCAAGCCTACTAGAGCAAGAAACCAGTTGTAATATAAAAT  
GCACTGCCCTACTGTTGGTATGACTACCGTTACCTACTGTTGTCATTGTTATTACAGCTATGG  
CCACTATTATTAAAGAGCTGTGTAACATCTCTGGCAAAAAAAAAAAAA

333/615

**FIGURE 330**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA68885
><subunit 1 of 1, 432 aa, 1 stop
><MW: 47644, pI: 5.18, NX(S/T): 2
MLQDPDSDQPLNSLDVKPLRKPRIPMETFRKVGIPIIIALSLASIIIVVVLIKVILDKYYFL
CGQPLHFIPRKQLCDGELDCPLGEDEEHCVKSFPEGPAVAVRLSKDRSTLQVLDSATGNWFSA
CFDNFTEALAETACRQMGYSRAVEIGPDQDLDVVEITENSQELMRNSSGPGCLSGSLVSLHCL
ACGKSLKTFRVVGEEASVDSWPWQVSIQYDKQHVC GGSILDPHWVLTAAHCFRKHTDVFNWK
VRAGSDKLGSFPSLAVAKIIIIIEFNPMYPKDNDIALMKLQFPLTFSGTVRPICLPFFDEELTP
ATPLWIIIGWGFTKQNGGKMSDILLQASVQVIDSTRCNADDAYQGEVTEKMMCAGIPEGGVDTC
QGDSGGPLMYQSDQWHVVGIVSWGYGCGGPSTPGVYTKVSAYLNWIYNVWKAEL
```

**Transmembrane domain:**

amino acids 32-53 (typeII)



334/615

**FIGURE 331**

AGTGGTTCGATGGGAAGGATCTTTCTCCAAGTGGTTCCTCTTGAGGGGAGCATTCTGCTGGC  
TCCAGGACTTTGGCCATCTATAAAGCTTGGCA**ATG**AGAAATAAGAAAATTCTCAAGGAGGACG  
AGCTCTTGAGTGAGACCCAACAAGCTGCTTTTCACCAAAATTGCAATGGAGCCTTTTGAAATCA  
ATGTTCCAAAGCCCAAGAGGAGAAATGGGGTGAACCTCTCCCTAGCTGTGGTGGTCATCTACC  
TGATCCTGCTCACCGCTGGCGCTGGGCTGCTGGTGGTCCAAGTTCTGAATCTGCAGGCGCGGC  
TCCGGGTCTTGAGATGTATTTCTCAATGACACTCTGGCGGCTGAGGACAGCCCGTCCTTCT  
CCTTGCTGCAGTCAGCACACCCTGGAGAACACCTGGCTCAGGGTGCATCGAGGCTGCAAGTCC  
TGCAGGCCCAACTCACCTGGGTCCGCGTCAGCCATGAGCACTTGCTGCAGCGGGTAGACAAC  
TCACTCAGAACCCAGGGATGTTTCAAGTCAAAGGTGAACAAGGCGCCCCAGGTCTTCAAGGTC  
ACAAGGGGGCCATGGGCATGCCTGGTGCCCCCTGGCCCGCCGGGACCACCTGCTGAGAAGGGAG  
CCAAGGGGGCTATGGGACGAGATGGAGCAACAGGCCCTCGGGACCCCAAGGCCACCGGGAG  
TCAAGGGAGAGGCGGGCCTCCAAGGACCCAGGGTGCTCCAGGGAAGCAAGGAGCCACTGGCA  
CCCCAGGACCCCAAGGAGAGAAGGGCAGCAAAGGCGATGGGGGTCTCATTTGGCCCCAAAGGGG  
AAACTGGAATAAGGGAGAGAAAGGAGACCTGGGTCTCCAGGAAGCAAAGGGGACAGGGGCA  
TGAAAGGAGATGCAGGGGTCTGGGGCCTCCTGGAGCCCAGGGGAGTAAAGGTGACTTCGGGA  
GGCCAGGCCACCAGGTTTGGCTGGTCTTCTGGAGCTAAAGGAGATCAAGGACAACCTGGAC  
TGCAGGGTGTTCCGGGCCCTCCTGGTGAGTGGGACACCCAGGTGCCAAGGGTGAGCCTGGCA  
GTGCTGGCTCCCCTGGGCGAGCAGGACTTCCAGGGAGCCCCGGGAGTCCAGGAGCCACAGGCC  
TGAAAGGAAGCAAAGGGGACACAGGACTTCAAGGACAGCAAGGAAGAAAAGGAGAATCAGGAG  
TTCCAGGCCCTGCAGGTGTGAAGGGAGAACAGGGGAGCCCAGGGCTGGCAGGTCCCAAGGGAG  
CCCCTGGACAAGCTGGCCAGAAGGGAGACCAGGGAGTGAAAGGATCTTCTGGGGAGCAAGGAG  
TAAAGGGAGAAAAAGGTGAAAGAGGTGAAACTCAGTGTCCGTGAGGATTGTCGGCAGTAGTA  
ACCGAGGCCGGGCTGAAGTTTACTACAGTGGTACCTGGGGGACAATTTGCGATGACGAGTGGC  
AAAATTCTGATGCCATTGTCTTCTGCCGATGCTGGGTACTCCAAAGGAAGGGCCCTGTACA  
AAGTGGGAGCTGGCACTGGGCAGATCTGGCTGGATAATGTTCAGTGTGCGGGCACGGAGAGTA  
CCCTGTGGAGCTGCACCAAGAATAGCTGGGGCCATCATGACTGCAGCCACGAGGAGGACGCAG  
GCGTGGAGTGCAGCGTCT**TGA**CCCGGAAACCCTTTCACTTCTCTGCTCCCGAGGTGTCCTCGGG  
CTCATATGTGGGAAGGCAGAGGATCTCTGAGGAGTCCCTGGGGACAACCTGAGCAGCCTCTGG  
AGAGGGGCCATTAATAAAGCTCAACATCATTGA

335/615

**FIGURE 332**

></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA68886  
><subunit 1 of 1, 520 aa, 1 stop  
><MW: 52658, pI: 9.16, NX(S/T): 3  
MRNKKILKEDELLSETQQAAFHQIAMEPFEINVPKPKRRNGVNFSLAVVVIYLLILLTAGAGLL  
VVQVLNLQARLRVLEMYFLNDTLAAEDSPSFSLLQSAHPGEHLAQGASRLQVLQAQLTWVRVS  
HEHLLQRVDNFTQNPGMFRIKGEQGAPGLQGHKGAMGMPGAPGPPGPPAEKGAKGAMGRDGAT  
GPSGPQGPPEGVKGEAGLQGPQGAPGKQGATGTPGPQGEKGSKGDGGLIGPKGETGTKGEKGD  
GLPGSKGDRGMKGDAGVMGPPGAQGSKGDFGRPGPPGLAGFPAGAKGDQGPGLQGVPPGAV  
GHPGAKGEPGSAGSPGRAGLPGSPGSPGATGLKGSKGDGTGLQGGQGRKGESGVPGPAGVKGEQ  
GSPGLAGPKGAPGQAGQKGDQGVKGSSGEQGVKGEKGERGENSVSVRIVGSSNRGRAEVYYSG  
TWGTICDDEWQNSDAIVFCRMLGYSKGRALYKVGAGTGQIWL DNVQCRGTESTLWSCTKNSWG  
HHDCSHEEDAGVECSV

**Transmembrane domain:**

amino acids 47-66 (type II)

**N-glycosylation sites.**

amino acids 43-47, 83-87, 136-140

**Tyrosine kinase phosphorylation site.**

amino acids 432-440

**N-myristoylation sites.**amino acids 41-47, 178-184, 253-259, 274-280, 340-346, 346-352,  
400-406, 441-447, 475-481, 490-496, 515-521**Amidation site.**

amino acids 360-364

**Leucine zipper pattern.**

amino acids 56-78

**Speract receptor repeat**

amino acids 422-471, 488-519

**Clq domain proteins.**

amino acids 151-184, 301-334, 316-349

336/615

**FIGURE 333**

GGGCTGTTGATTTGTGGGGGATTTTGAAGAGAGGAGGAATAGGAGGAAGGGGTTGAGGGGCTG  
CCTCTGGCATATGCACACACTCACACATTCTGTACACCCGTCACACACACATACCATGTTCT  
CCATCCCCCAGGTCCAGCCCTCAGTGCTGTCCCATCCAGCAGGGCTACCCTGAAGCTCTGGC  
TGCAGCCCTCCCGTCCAGTGGGCAGGCGGCTTCATCCCTCCTTTCTCTCCCAAAGCCCAACTG  
CTGTCACTGCATGCTCTGCCAAGGAGGAGGGAAGTGCAGTGACAGCAGGAGTAAGAGTGGGAG  
GCAGGACAGAGCTGGGACACAGGTATGGAGAGGGGGTTGAGCGAGCCTAGAGAGGGCAGACTA  
TCAGGGTGCCGGCGGTGAGAATCCAGGGAGAGGAGCGGAAACAGAAGAGGGGCAGAAGACCGG  
GGCACTTGTGGGTTGCAGAGCCCCTCAGCC**ATG**TTGGGAGCCAAGCCACACTGGCTACCAGGT  
CCCCTACACAGTCCCGGGCTGCCCTTGCTTCTGGTGGTCTTGGCCCTGGGGGCCGGGTGGGCC  
CAGGAGGGGTGAGAGCCCGTCCCTGCTGGAGGGGGAGTGCCTGGTGGTCTGTGAGCCTGGCCGA  
GCTGCTGCAGGGGGGCCCGGGGAGCAGCCCTGGGAGAGGCACCCCTGGGCGAGTGGCATT  
GCTGCGGTCCGAAGCCACCACCATGAGCCAGCAGGGGAAACCGCAATGGCACCAGTGGGGCC  
ATCTACTTCGACCAGGTCTTGGTGAACGAGGGCGGTGGCTTTGACCGGGCCTCTGGCTCCTTC  
GTAGCCCCGTGCCGGGTGTCTACAGCTTCCGGTTCCATGTGGTGAAGGTGTACAACCGCCAA  
ACTGTCCAGGTGAGCCTGATGCTGAACACGTGGCCTGTCTCATCTCAGCCTTTGCCAATGATCCT  
GACGTGACCCGGGAGGCAGCCACCAGCTCTGTGCTACTGCCCTTGGACCCTGGGGACCGAGTG  
TCTCTGCGCTGCGTCGGGGGAATCTACTGGGTGGTGGAAATACTCAAGTTTCTCTGGCTTC  
CTCATCTTCCCTCTCT**TG**AGGACCCAAGTCTTTCAAGCACAGAATCCAGCCCCTGACAACTTT  
CTTCTGCCCTCTCTTGCCCCAGAAACAGCAGAGGCAGGAGAGAGACTCCCTCTGGCTCCTATC  
CCACCTCTTTGCATGGGACCCTGTGCCAAACACCCAAGTTTAAGAGAAGAGTAGAGCTGTGGC  
ATCTCCAGACCAGGCCTTTCCACCCACCCACCCCAAGTTACCCTCCCAGCCACCTGCTGCATC  
TGTTCTCTGCTGCTCCTCCGGTTCCCCCACCCAGCTTCTGCTCAATGCTGATCAGGGACAGG  
TGGCGCAGGTGAGCCTGACAGGCCCCACAGGAGCCAGATGGACAAGCCTCAGCGTACCCTG  
CAGGCTTCTTCTGTGAGGAAAGCCAGCATCACGGATCTCAGCCAGCACCGTCAGAAGCTGAG  
CCAGCACCGTATGGGCTAGGGTGGGAGGCTCAGCCACAGGCAGAAGGGTGGGAAGGGCCTGGA  
GTCTGTGGCTGGTGAAGGAAGGAGGGTGTATTGTCTAGACTGAACATGGTACACATTCTG  
CATGTATAGCAGAGCAGCCAGCAGGTAGCAATCCTGGCTGTCTTCTATGCTGGATCCCAGAT  
GGA CTCTGGCCCTTACCTCCCCACCTGAGATTAGGGTGAGTGTGTTTGCTCTGGCTGAGAGCA  
GAGCTGAGAGCAGGTATACAGAGCTGGAAGTGGACCATGGAAAACATCGATAACCATGCATCC  
TCTTGCTTGGCCACCTCCTGAAACTGCTCCACCTTTGAAGTTTGAACCTTTAGTCCCTCCACAC  
TCTGACTGCTGCCTCCTTCCCTCCAGCTCTCTCACTGAGTTATCTTCACTGTACCTGTTCCAG  
CATATCCCCACTATCTCTCTTCTCCTGATCTGTGCTGTCTTATTCTCCTCCTTAGGCTTCT  
ATTACCTGGGATTCCATGATTCACTTCTTCAGACCTCTCCTGCCAGTATGCTAAACCTCCC  
TCTCTCTTTCTTATCCCGCTGTCCCATTTGGCCAGCCTGGATGAATCTATCAATAAAACAACT  
AGAGAATGGTGGTCAGTGAGACACTATAGAATTACTAAGGAGAAGATGCCTCTGGAGTTTGA  
TCGGGTGTTACAGGTACAAGTAGGTATGTTGCAGAGGAAATAAATATCAAACTGTATACTAA  
AATTAAAA

337/615

**FIGURE 334**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71180
><subunit 1 of 1, 205 aa, 1 stop
><MW: 21521, pI: 7.07, NX(S/T): 1
MLGAKPHWLPGLPLVLVLLALGAGWAQEGSEPVLLERGECLVVCEPGRAAAGGPGGAA
LGEAPPGRVAFAAVRSHHHEPAGETGNGTSGAIYFDQVLVNEGGGFDRASGSFVAPVRGVYSF
RFHVVKVYNRQTVQVSLMLNTWPVISAFANDPDVTREAAATSSVLLPLDPGDRVSLRLRRGNLL
GGWKYSSFSGFLIFPL
```

**Signal peptide:**  
amino acids 1-32

338/615

**FIGURE 335**

[illegible]

339/615

**FIGURE 336**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71184
><subunit 1 of 1, 388 aa, 1 stop
><MW: 43831, pI: 9.64, NX(S/T): 3
MKTLLIAAYSGVLRGERQAEADRSQRSHGGPALSREGSGRWGTGSSILSALQDLFSVTWLNRSK
VEKQLQVISVLQWVLSFLVLGVACSAILMYIFCTDCWLI AVL YFTWLVFDWNTPKKGRRSQW
VRNWAVWRYFRDYFPIQLVKTHNLLTTRNYIFGYHPHGIMGLGAF CNFSTEATEVSKKFPGIR
PYLATLAGNFRMPVLREYLMSGGICPVSRDTIDYLLSKNGSGNAIIIVVGGAESLSSMPGKN
AVTLRNRKGFVKLALRHGADLVPIYSFGENEVYKQVIFEEGSWGRWVQKFKQYIGFAPCI FH
GRGLFSSDTWGLVPYSKPITTVVGEPITIPKLEHPTQQDIDLYHTMYMEALVKLFDKHKTKEG
LPETEVLEV N
```

**Important features of the protein:****Transmembrane domain:**

amino acids 76-97

**N-glycosylation sites.**

amino acids 60-63, 173-176, 228-231

**N-myristoylation sites.**

amino acids 10-15, 41-46, 84-89, 120-125, 169-174, 229-234, 240-245, 318-323, 378-383

340/615

**FIGURE 337**

GGGCGGCGGGATGGGGGCCGGGGGCGGGGCGCGCACTCGCTGAGGCCCCGACGCAGGGCCGGGCGGGCCCA  
GGGCGGAGGAGCGCGGCGGCCAGAGCGGGGCGCGGAGGCGACGCCGGGACGCCCCGCGACGAGCAGGTGGCG  
GCGGCTGCAGGCTTGTCCAGCCGAAGCCCTGAGGGCAGCTGTTCCCACTGGCTCTGCTGACCTTGTGCCTTGA  
CGGCTGTCTCAGCGAGGGGCGTGCACCCGCTCCTGAGCAGCGCCATGGGCCTGCTGGCCTTCTGAAGACCCA  
GTTCTGTCTGCACCTGCTGGTGGCTTTGTCTTCGTGGTGGTGGTCTGGTTCATCAACTTCGTCCAGCTGTGCAC  
GCTGGCGCTCTGGCCGGTCAGCAAGCAGCTCTACCGCCGCTCAACTGCCGCTCGCTACTCACTCTGGAGCCA  
ACTGGTCATGCTGCTGGAGTGGTGGTCTGACCGAGGTGTACACTGTTACGGACCAGGCCACGGTAGAGCGCTT  
TGGGAAGGAGCACGCAGTCATCATCTCAACCACAACCTTCGAGATCGACTTCTCTGTGGGTGGACCATGTGTGA  
GCGCTTCGGAGTGTGGGGAGCTCCAAGGTCTCGCTAAGAAGGAGCTGCTCTACGTGCCCTCATCGCTGGAC  
GTGGTACTTTCTGGAGATTGTGTTCTGCAAGCGGAAGTGGGAGGAGGACCGGGACACCGTGGTGAAGGGCTGAG  
GCGCTGTGCGACTACCCCGAGTACATGTGGTTCTCCTGTACTGCGAGGGGACCGCTTCACGGAGACCAAGCA  
CCGCTTAGCATGGAGGTGGCGGCTGCTAAGGGGCTTCTGTCTCAAGTACCACCTGCTGCCGCGGACCAAGGG  
CTTCACCACCGCAGTCAAGTGCCTCCGGGGGACAGTCGAGCTGTCTATGATGTAACCTGAACCTCAGAGGAAA  
CAAGAACCCGCTCCCTGCTGGGGATCCTCTACGGGAAGAAGTACGAGCGGACATGTGCGTGAGGAGATTTCTCT  
GGAAGACATCCCGCTGGATGAAAAGGAAGCAGCTCAGTGGCTTCATAAACTGTACCAGGAGAAGGACGCGCTCCA  
GGAGATATATAATCAGAAGGGCATGTTTCCAGGGGAGCAGTTTAAAGCCTGCCCGGAGGCGCTGGACCTCCTGAA  
CTTCCTGTCTGGGCCACCATTTCTCTGTCTCCCTCTTCAGTTTTGTCTTGGGCGCTTTGCCAGCGGATCACC  
TCTCTGATCCTGACTTTCTTGGGGTTTGTGGGAGCAGCTTCTTTGGAGTTCGCAGACTGATAGGAGAATCGCT  
TGAACCTGGGAGGTGGAGATTGCAGTGAAGCTGAGATGGCATCACTGTACTCCAGCCTAGGCAACAGAGCAAGACT  
CAGTCTCAAAAAAAAAAAAAAAAAACAAAAACCCAGAAATCTGGAGTTGAAGTGTGTAGTTACTGACATGAAAA  
ATTCAGTAGAGGCTGAACAGCAGATTTGAGCAGGCAGAAAAAATCAGCAAGCTTGAAGATGGTACCTTGAGATT  
TTTCAGGCTAATGAAAAAGAATGAAGGAAAATTAACAGCCTCAGAGACCCATGGTGCACCGTCACACAAATCAA  
CATATGCATGATGAGAGTCCAGAGGAGAGAGAGAAAGGTCAGAAAGAATGGCCACAAGCTGATGAAAAACA  
GTAACCTACCCACTCAGGAAGCTCAGTGAACCTCAATGAGGATGAATATCAGAGATCCACACCTAGATATTTTCA  
AATCAAAGTGTCAAATGACAAAGAATCTTGAAAGCAGCAAGAGATGAGCAACTTATCTTGTCAAAGGATCTTG  
ATCAGATTAACAGCTCATTTCTCTCAGAAATCATGGGAGCCAGGAGATAGTGGGATGAACACTGTTGAAGGCAA  
AACCTTCAACTGTAATTATTGGACTTTTGAGTCTTAGATGGTCTGACCTCTTGTCTTCAGGGACAGTTTTTCA  
ATTTAATCCCTAATAACAATTAGTCAAGCTTCTTGACCTGTAGGAAGGCCTGTCTTTAGGCCGGGCACAGTGGC  
TTACACCTGTAATCCAGCACTTTGGGAGGCCAGACGGGTGGATCATTTGGGGTCAGGCTGATCTCAAACCTCCT  
GAGTTCAGGTGATCTGCCCGCTCAGCCTCCCAAAGTGTGTGATTGCAGGCGTGAGCCACTGCGCCTGGCCGGA  
ATTTCTTTTAAAGGCTGAATGATGGGGGCCAGGCACGATGGCTCACGCCTGTGATCCCAAGTAGCTTGATTGTA  
AACATGCACCAACCATGCCTGGCTAATTTTTGTATTTTAGTAGAGACGTGTTAGCCAGGCTGGTCTCGATCTCCT  
GACCTCAAGTGACCACCTGCCTCAGCCTCCCAAAGTACTGGGATTACAGGCGTGAGCCACTGTGCCTGGCCTTGA  
GCATCTTGTGATGTGCTTATTGGCCATTTGTATATCTTCTATCTTCTTTGGGAAATGCTGTTCAGTCTCTTG  
CCTTTTTTAAATTTTTATTATTTATTTATTTATTTATTTTGGAGACAGGCTTGTGTTCTGTTGCCAGGCTGGAGTA  
CAGTGGCACAGTCTTGGCTCACTGCAGCCTCGACCTCCTGGGCTGCAGTGATCCTCCACCTCAGCCTCCCTTGT  
AGCTGTATTTTTTGTATTTTGTATTTGTAGCTGTAGTTTTTGTATTTTTTGTGGAGACAGATTTACCATGA  
TGCCAGGCTGGTCTTGAACCTCTGAGCTCAAGTGATCTGCCTGCTTCAGCCTCCCAAAGTGCTGGGATTACAGA  
CATGAGCCACTGCACCTGGCAAACCTCCCAAATTCACACACACACACAAAAACCACCTGATTCAAATGGGCA  
GAGGGGCGGGTGTGGCCCCAACTACCAGGGAGACTGAAGTGGGAGGATCGCTTGGGCATGAGAAGTCGAGGCTG  
CAGTGAGTCGAGGTTGTGCGACTGCATTCCAGCCTGGACAACAGAGTGAGACCCTGTCTC

341/615

**FIGURE 338**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71213
><subunit 1 of 1, 368 aa, 1 stop
><MW: 42550, pI: 9.11, NX(S/T): 1
MGLLAFLKTQFVLHLLVGFFVVSGLVINQVQLCTLALWPVSKQLYRRRLNCRLAYSLWSQLVM
LLEWWSCTECTLFQATVERFGKEHAVIILNHNFEIDFLCGWTMCERFGVLGSSKVLAKKEL
LYVPLIGWTWYFLEIVFCKRKWEEDRDTVVEGLRRLSDYPEYMWFLLYCEGTRFTETKHRVSM
EVAAAKGLPVLKYHLLPRTKGFTTAVKCLRGTVAAVDVTLNFRGNKNPSLLGILYGKKYEAD
MCVRRFPLEDIPLDEKEAAQWLHKLYQEKDALQEIYNQKGMFPGEQFKPARRPWTLNFLSWA
TILLSPLFSFVLGVFASGSPLLILTF LGFVGAASFGVRRRLIGESLEPGRWRLQ
```

**Important features of the protein:****Signal peptide:**

amino acids 1-25

**Transmembrane domains:**

amino acids 307-323, 335-352

**Tyrosine kinase phosphorylation sites.**

amino acids 160-168, 161-169



342/615

**FIGURE 339**

GATATTCTTTATTTTAAAGAATCTGAAGTACTATGCATCACTCCCTCCAATGTCCTGGGGCAG  
CCACCAGGCATATTCATCTTTGTGTGTGTTTTCTTTGCTTTAGCACTGGGGCACTTCTTGC  
TTATTTCTTTGGTAGGAAAGGGGCTCAGTTTGTCTTGTGGGGTTGGTGGCAGGCAGGCCGGCT  
TACGCCTGATACGGCCCTGGGTTAGAAGGGAAGGGAAGATAAACTTTTATACAAATGGGGATA  
GCTGGGGTCTGAGACCTGCTTCCTCAGTAAAATTCTTGGGATCTGCCTATACCTTCTTTTCTC  
TAACCTGGCATACCTGCTTAAAGCCTCTCAGGGCTTCTCTCTGTTCTTAGGATCAAAGTATT  
TAGAGCTACAAGAGCCCTCATGGTCTGGCCCCCTGCCCCCTGGCCAGCTTCATTGTACATGTG  
GTGTTCTCTTGTCTTCCTGTAATGTGGTATGCCATGGGGTCTTGCACAAGCCTTTCCTCTT  
TGGCTGGACACTGTTCCCTGCCCCCCCCATACTCTTCCTACTTAATATGTAGTCATCCTGCAG  
ATTTCAATTCTAACATCATTTTCTCCAGGGATCCTGGCCTGACAGAATCTCATCTTGTTTAAT  
GCTCTCATAAGACCACTTGTTTCCCTTTTGCAGCACTTGCCACTCAGTTGTATCTTTATGTGC  
GTTTGTGGTTGTATGGGTTGTGTCTGTTCCCCAGAATGCCAGCTCTGAGCTGCGTGAGGGTC  
AAGGGCATTGCTGTGCCTGCCAGGTATAGTGCCTACATGTGGTGGGTGCTCATGTTTTAGAGA  
CTAAATGGAGGAGGAGATGAGGAAAAGATTGAAATCTCTCAGTTCACCAGATGGTGTAGGGCC  
CAGCATTGTAAATTCACACGTTGACTGTGCTTGTGAATTATCTGGGGATGCAGGTCCCTGATTC  
AGTAGGCCCAGGTTGGGCATCTCTAACAACTCCACGCTGATGCTGATGCTGGTCCTATGAAC  
TATACTAAATAGTAAGAATCTATGGAGCCAGGCTGGGCATGGTGGCTCACACCTATGATCCCA  
GCACTTTGGGAGGCTGAGGCAGGCTGATCACCTGGAGTCAGGATTTCAAGACTAGCCTGGCCA  
ACATGGTGGAAACCCCATCTGTACTAAAAATACACAAATTAGCTGGGCATGGTGGCACATGCCT  
GTAGTCCCAGCTACTTGGGAGGCTGAAGCAAGAGAATCGCTTGAACCTGGGAGGCGGAGGTTG  
CAGTGAGCCGAGATCAGGCCACTGTATTCCAACCAGGGTGACAGAGTGAGACTCTATGTCCAA  
AAAAAAAAAA

343/615

**FIGURE 340**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71234  
><subunit 1 of 1, 143 aa, 1 stop  
><MW: 15624, pI: 9.58, NX(S/T): 0  
MHHSLQCPGAATRHHLCVCFSEFALALGHFLLISLVGKGLSLSCGVGGRQAGLRLIRPWVRE  
GKINFYTNNGDSWGLRPASSVKFLGSAYTFFSLTWHTLLKASQGFSFLGSKYLELQEPSWSGP  
CPPGQLHCTCGVLLSFL

**Important features of the protein:**

**Signal peptide:**

amino acids 1-28

344/615

**FIGURE 341**

CGCCATGGCCGGGCTATCCCGCGGGTCCGCGCGCGCACTGCTCGCCGCCCTGCTGGCGTCGACG  
CTGTTGGCGCTGCTCGTGTGCGCCGCGCGGGGTGCGGGCGGGCCGGGACCACGGGGACTGGGAC  
GAGGCCTCCCGGCTGCCCGCGCTACCACCCCGCGAGGACGCGGCGCGCGTGGCCCGCTTCGTG  
ACGCACGTCTCCGACTGGGGCGCTCTGGCCACCATCTCCACGCTGGAGGCGGTGCGCGGCCGG  
CCCTTCGCCGACGTCTCTCGCTCAGCGACGGGCCCCCGGGCGCGGGCAGCGGCGTGCCCTAT  
TTCTACCTGAGCCCGCTGCAGCTCTCCGTGAGCAACCTGCAGGAGAATCCATATGCTACACTG  
ACCATGACTTTGGCACAGACCAACTTCTGCAAGAAACATGGATTGATCCACAAAGTCCCCTT  
TGTGTTACATAATGCTGTCAGGAAGTGTGACCAAGGTGAATGAAACAGAAATGGATATTGCA  
AAGCATTGCTTATTCATTCGACACCCTGAGATGAAAACCTGGCCTTCCAGCCATAATTGGTTC  
TTTGCTAAGTTGAATATAACCAATATCTGGGTCCTGGACTACTTTGGTGGACCAAAAATCGTG  
ACACCAGAAGAATATTATAATGTCACAGTTCAGTGAAGCAGACTGTGGTGAATTTAGCAACAC  
TTATGAAGTTTCTTAAAGTGGCTCATACACACTTAAAAGGCTTAATGTTTCTCTGGAAAGCGT  
CCCAGAATATTAGCCAGTTTTCTGTC

345/615

**FIGURE 342**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71269
><subunit 1 of 1, 220 aa, 1 stop
><MW: 24075, pI: 7.67, NX(S/T): 3
MAGLSRGSARALLAALLASTLLALLVSPARGRGGRDHGDWDEASRLPPLPPREDAARVAR
FVTHVSDWGALATISTLEAVRGRPFADVLSLSDGPPGAGSGVPYFYLSPLQLSVSNLQEN
PYATLTMTLAQTNFCKKHGFD PQSPLCVHIMLSGTVTKVNETEMDIAKHSLFIRHPMKT
WPSSHNWFFAKLNITNIWVLDYFGGPKIVTPEEYYNVTVQ
```

**Important features of the protein:****Transmembrane domain:**

Amino acids 11-29

**N-glycosylation sites:**

Amino acids 160-164;193-197;216-220

**N-myristoylation sites:**

Amino acids 3-9;7-13;69-75;97-103

346/615

**FIGURE 343**

GGCTGGACTGGAACCTCCTGGTCCCAAGTGATCCACCCGCCTCAGCCTCCCAAGGTGCTGTGAT  
TATAGGTGTAAGCCACCGTGTCTGGCCTCTGAACAACCTTTTTTCAGCAACTAAAAAAGCCACAG  
GAGTTGAACTGCTAGGATTCTGACTATGCTGTGGTGGCTAGTGCTCCTACTCCTACCTACATT  
AAAATCTGTTTTTTGTTCTCTTGTAAGTACGCTTTACCTTCCTAACACAGAGGATCTGTCACT  
GTGGCTCTGGCCCCAACCTGACCTTCACCTCTGGAACGAGAACAGAGGTTTCTACCCACACCGT  
CCCCTCGAAGCCGGGGACAGCCTCACCTTGCTGGCCTCTCGCTGGAGCAGTGCCCTCACCAAC  
TGTCTCACGTCTGGAGGCACTGACTCGGGCAGTGCAAGGTAGCTGAGCCTCTTGGTAGCTGCGG  
CTTTCAAGGTGGGCCTTGCCCTGGCCGTAGAAGGGATTTGACAAGCCCGAAGATTTCATAGGCG  
ATGGCTCCCACTGCCCAGGCATCAGCCTTGCTGTAGTCAATCACTGCCCTGGGGCCAGGACGG  
GCCGTGGACACCTGCTCAGAAGCAGTGGGTGAGACATCACGCTGCCCGCCCATCTAACCTTTT  
CATGTCCTGCACATCACCTGATCCATGGGCTAATCTGAACTCTGTCCCAAGGAACCCAGAGCT  
TGAGTGAGCTGTGGCTCAGACCCAGAAGGGGTCTGCTTAGACCACCTGGTTTATGTGACAGGA  
CTTGCAATTCTCCTGGAACATGAGGGAACGCCGGAGGAAAGCAAAGTGGCAGGGAAGGAACTTG  
TGCCAAATTATGGGTCAGAAAAGATGGAGGTGTTGGGTATCACAAGGCATCGAGTCTCCTGC  
ATTCAAGTGGACATGTGGGGGAAGGGCTGCCGATGGCGCATGACACACTCGGGACTCACCTCTG  
GGGCCATCAGACAGCCGTTTCCGCCCCGATCCACGTACCAGCTGCTGAAGGGCAACTGCAGGC  
CGATGCTCTCATCAGCCAGGCAGCAGCCAAAATCTGCGATCACAGCCAGGGGCAGCCGCTCTG  
GGAAGGAGCAAGCAAAGTGACCATTTCTCCTCCCCCTCCTTCCCTCTGAGAGGGCCCTCCTATGT  
CCCTACTAAAGCCACCAGCAAGACATAGCTGACAGGGGGCTAATGGCTCAGTGTTGGCCCAGGA  
GGTCAGCAAGGCCTGAGAGCTGATCAGAAGGGCCTGCTGTGCGAACACGGAAATGCCTCCAGT  
AAGCACAGGCTGCAAAATCCCCAGGCAAAGGACTGTGTGGCTCAATTTAAATCATGTTCTAGT  
AATTGGAGCTGTCCCCAAGACCAAAGGAGCTAGAGCTTGGTTCAAATGATCTCCAAGGGCCCT  
TATACCCCAGGAGACTTTGATTTGAATTTGAAACCCCAAATCCAAACCTAAGAACCAGGTGCA  
TTAAGAATCAGTTATTGCCGGGTGTGGTGGCCTGTAATGCCAACATTTTGGGAGGCCGAGGCG  
GGTAGATCACCTGAGGTCAGGAGTTCAAGACCAGCCTGGCCAACATGGTGAAACCCCTGTCTC  
TACTAAAAATACAAAAAACTAGCCAGGCATGGTGGTGTGTGCCTGTATCCCAGCTACTCGGG  
AGGCTGAGACAGGAGAATTACTTGAACCTGGGAGGTGAAGGAGGCTGAGACAGGAGAATCACT  
TCAGCCTGAGCAACACAGCGAGACTCTGTCTCAGAAAAAATAAAAAAGAATTATGGTTATTT  
GTAA

347/615

**FIGURE 344**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71277
><subunit 1 of 1, 109 aa, 1 stop
><MW: 11822, pI: 8.63, NX(S/T): 0
MLWWLVLLLLPTLKSVFCSLVTSLYLPNTEDLSLWLWPKPDLHSGTRTEVSTHTVPSKPGTAS
PCWPLAGAVPSPTVSRLEALTRAVQVAEPLGSCGFQGGPCPGRRRD
```

**Signal peptide:**  
amino acids 1-15

348/615

**FIGURE 345**

[illegible]

349/615

**FIGURE 346**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71286

&gt;&lt;subunit 1 of 1, 671 aa, 1 stop

&gt;&lt;MW: 74317, pI: 7.61; NX(S/T): 0

MPHAFKPGDLVFAKMKGYPHWPARIDDIADGAVKPPPKNKYPIFFFGTHETAFLGPKDLFPYDK  
CKDKYGKPNKRKGFNEGLWEIQNNPHASYSAPPPVSSSDSEAPEANPADGSDADEDDERGV  
AVTAVTATAASDRMESDSDSDKSSDNSGLKRKTPALKMSVSKRARKASSDLDQASVSPSEEN  
SESSSESEKTSQDFTPEKKA AVRAPRRGPLGGRKKKKAPSASDSDSKADSDGAKPEPVAMAR  
SASSSSSSSSSSSDSDSVKKPPRGRKPAEKPLPKPRGRKPKPERPPSSSSSDSDSDEVDRISE  
WKRRDEARRRELEARRRREQEEELRRLREQEKEEKERRRERADRGEAERGSGGSSGDELREDD  
EPVKKRGRKGRGRGPPSSSDSEPEAELEAREAKKSAKKPQSSSTEPARKPGQKEKRV RPEEKQQ  
AKPVKVERTRKRSEGFMDRKVEKKKEPSVEEKLQKLHSEIKFALKVDS PDVKRCLNALEELG  
TLQVTSQILQKNTDVVATLKKIRRYKANKDVMKAAEVYTRLKSRVLGPKIEAVQKVNKAGME  
KEKAEKLAGEEELAGEEAPQEKAEDKPDLSAPVNGEATSQKGESAEDKEHEEGRDSEEGPR  
CGSSEDLHDSVREGPDLDRPGSDRQERERARGDSEALDEES

**Signal peptide:**

amino acids 1-13



350/615

**FIGURE 347**

GTTGGTTCTCCTGGATCTTCACCTTACCAACTGCAGATCTTGGGACTCATCAGCCTCAATAATTATATTAATTA  
ACACCATTGTGAAGAGAACATTGTTTTTCATCATGAATGCTAATAAAGATGAAAGACTTAAAGCCAGAAGCCAAGA  
TTTTACCTTTTTCTGCTTTGATGATGCTAAGCATGACCATGTTGTTTCTTCCAGTCACTGGCAGCTTTGAAGCA  
AAATATTCCAAGACTCAAGCTAACCTACAAAGACTTGCTGCTTTCAAATAGCTGTATTCCCTTTTTGGGTTTCATC  
AGAAGGACTGGATTTTCAAACCTCTCTCTTAGATGAGGAAAGAGGCAGGCTGCTCTTGGGAGCCAAAGACCACAT  
CTTTCTACTCAGTCTGGTTGACTTAAACAAAATTTTAAAGAAGATTTATTGGCCTGCTGCAAAGGAACGGGTGGA  
ATTATGTAAATTAGCTGGGAAAGATGCCAATACAGAATGTGCAAATTCATCAGAGTACTTCAGCCCTATAACAA  
AACTCACATATATGTGTGTGGAACCTGGAGCATTTCATCCAATATGTGGGTATATTGATCTTGGAGTCTACAAGGA  
GGATATTATATTCAAACCTAGACACACATAATTTGGAGTCTGGCAGACTGAAATGTCCTTTTCGATCCTCAGCAGCC  
TTTTGCTTCAGTAATGACAGATGAGTACCTCTACTCTGGAACAGCTTCTGATTTTCCTTGGCAAAGATACTGCATT  
CACTCGATCCCTTGGGCCTACTCATGACCACCACTACATCAGAAGTACATTTTCAGAGCACTACTGGCTCAATGG  
AGCAAAATTTATTGGAACCTTTCTTCATACCAGACACCTACAATCCAGATGATGATAAAATATATTTCTTCTTTCTG  
TGAATCATCTCAAGAAGGCAGTACCTCCGATAAAACCATCCTTTCTCGAGTTGGAAGAGTTTGTAAGAATGATGT  
AGGAGGACAACGCAGCCTGATAAAACAGTGGACGACTTTTCTTAAGGCCAGACTGATTTGCTCAATTCCTGGAG  
TGATGGGGCAGATACTTACTTTGATGAGCTTCAAGATATTTATTACTCCCCACAAGAGATGAAAGAAATCCTGT  
AGTATATGGAGTCTTTACTACAACCAGCTCCATCTTCAAAGGCTCTGCTGTTTGTGTGTATAGCATGGCTGACAT  
CAGAGCAGTTTTTAATGGTCCATATGCTCATAAGGAAAGTGCAGACCATCGTTGGGTGCAGTATGATGGGAGAAT  
TCCTTATCCACGGCCTGGTACATGTCCAAGCAAACCTATGACCCACTGATTAAGTCCACCCGAGATTTTCCAGA  
TGATGTCATCAGTTTCATAAAGCGGCACTCTGTGATGTATAAGTCCGTATACCCAGTTGCAGGAGGACCACGTT  
CAAGAGAATCAATGTGGATTACAGACTGACACAGATAGTGGTGGATCATGTTCATTGCAGAAGATGGCCAGTACGA  
TGTAATGTTTCTTGGAACAGACATTGGAACCTGTCTCAAAGTTGTGTCAGCATTTCAAAGGAAAAGTGGAAATATGGA  
AGAGGTAGTCTGGAGGAGTTGCAGATATTCAAGCACTCATCAATCATCTTGAACATGGAAATGTCTCTGAAGCA  
GCAACAATTGTACATTGGTTCCCGAGATGGATTAGTTTCAGCTCTCCTTGCACAGATGCGACACTTATGGGAAAGC  
TTGCGCAGACTGTTGTCTTGCCAGAGACCCCTACTGTGCTTGGGATGGAATGCATGCTCTCGATATGCTCCTAC  
TTCTAAAAGGAGAGCTAGACGCCAAGATGTAAATATGGCGACCCCAATCACCCAGTGTGGGACATCGAAGACAG  
CATTAGTCATGAACTGCTGATGAAAAGGTGATTTTTTGGCATTGAATTTAACTCAACCTTTCTGGAATGTATACC  
TAAATCCCAACAAGCAACTATTAATGGTATATCCAGAGGTGAGGGGATGAGCATCGAGAGGAGTTGAAGCCCGA  
TGAAAGAATCATCAAAACGGAATATGGGCTACTGATTTCGAAGTTTGCAGAAGAAGGATTCTGGGATGTATTACTG  
CAAAGCCCAGGAGCACACTTTTCATCCACACCATAGTGAAGCTGACTTTGAATGTCATTGAGAATGAACAGATGGA  
AAATACCCAGAGGGCAGAGCATGAGGAGGGGCAGGTCAAGGATCTATTGGCTGAGTCACGGTTGAGATACAAAGA  
CTACATCCAAATCCTTAGCAGCCCCAACTTCAGCCTCGACCACTGCTGCGAACAGATGTGGCACAGGGAGAAGCG  
GAGACAGAGAAACAAGGGGGGCCAAAGTGAAGCACATGCAGGAAATGAAGAAGAAACGAAATCGAAGACATCA  
CAGAGACCTGGATGAGCTCCCTAGAGCTGTAGCCACGTAGTTTTCTACTTAATTTAAAGAAAAGAATTCCTTACC  
TATAAAAACCTATTGCCTTCTGTTTTGTATATCCCTTATGATTAATTCATAAATGCTTCCCATGGAGTTTTGCTAAGG  
CACAAGACAATAATCTGAATAAGACAATATGTGATTAATATAAGAAGGGGCAAAAATTCATTTGAACCATTTTT  
CCAAGAACAATCTGCACAAGCAAAGTATAAGAATTATCCTAAAAATAGGGGGTTTTACAGTTGTAAATGTTTTTA  
TGTTTTGAGTTTTTGGAAATTTATTGTTCATGTTAAATAGTTGAGCTAAGCAAGCCCCGAATTTGATAGTGTATAAGGT  
GCTTTATTCCCTCGAATGTCCATTAAGCATGGAATTTACCATGCAGTTGTGCTATGTTCTTATGAACAGATATAT  
CATTCTATTGAGAACCAGTACCTTGTGGTAGGGAATAAGAGGTCAGACACAAATTAAGACAACCTCCCATATC  
AACAGGAACCTTTCTCAGTGAGCCATTCACTCCTGGAGAATGGTATAGGAATTTGGAGAGGTGCATTATTTCTTTT  
TGGCCACTGGGGTTAAATTTAGTGTACTACAACATTGATTTACTGAAGGGCACTAATGTTTTCCCCCAGGATTTCT  
ATTGACTAGTCAGGAGTAACAGGTTTACAGAGAGAAGTTGGTGCTTAGTTATGTGTTTTTTAGAGTATATACTAA  
GCTCTACAGGGACAGAATGCTTAATAAATACTTTAATAAGATATGGGAAAATATTTAATAAAACAAGGAAAACA  
TAATGATGTATAATGCATCCTGATGGGAAGGCATGCAGATGGGATTTGTTAGAAGACAGAAGGAAAGACAGCCAT  
AAATCTGGCTTTGGGAAAACCTCATATCCCCATGAAAAGGAAGAACAATCACAAATAAAGTGAGAGTAATGTAA  
TGGAGCTCTTTTCACTAGGGTATAAGTAGCTGCCAATTTGTAATTCATCTGTTAAAAAAATCTAGATTATAACA  
AAGTCTAGCAAAATCTGAGGAAACATAAATCTTCTGGAAGAATCATAGGAAGAGTAGACATTTTATTATAACC  
AATGATATTTTCAGTATATATTTTCTCTTTTTAAAAAATATTTATCATACTCTGTATATTATTTCTTTTTACTGC  
CTTTATTCTCTCCTGTATATTGGATTTTGTGATTATATTGAGTGAATAGGAGAAAACAATATATAACACACAGA  
GAATTAAGAAAATGACATTTCTGGGGAGTGGGGATATATATTTGTTGAATAACAGAACGAGTGTAAAATTTTAAC  
AACGGAAGGGTTAAATTAACCTTTTGACATCTTCACTCAACCTTTTCTCATTGCTGAGTTAATCTGTTGTAATT  
GTAGTATTGTTTTGTAATTTAACAATAAATAAGCCTGCTACATGT

351/615

**FIGURE 348**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71883

&gt;&lt;subunit 1 of 1, 777 aa, 1 stop

&gt;&lt;MW: 89651, pI: 7.97, NX(S/T): 3

MNANKDERLKARSQDFHLFPALMMLSMTMLFLPVTGTLKQNIPLRLKLTYSKDLLLSNSCIPFLG  
SSEGLDFQTLTLLDEERGRLLLGAKDHIFFLLSLVDLNKNFKKIYWPAAKERVELCKLAGKDANT  
ECANFIRVLQPYNKTHIYVCGTGAFHPICGYIDLGVYKEDIIFKLDTHNLESGRLKCPFDPPQ  
PFASVMTDEYLYSGTASDFLGKDTAFTRSLGPTHDDHHYIRTDISEHYWLNKAKFIGTFFIPDT  
YNPDDDKIYFFRFRESSQEGSTSDKTILSRVGRVCKNDVGGQRS LINKWTTFLKARLICSIPGS  
DGADTYFDELQDIYLLPTRDERNPVVYGVFTTTSSIFKGSACVYSMA DIRAVFNGPYAHKES  
ADHRWVQYDGRIPYPRPGTCPSKTYDPLIKSTRDFPDVISFIKRHSVMYKSVYPVAGGPTEK  
RINV DYRLTQIVVDHVIAEDGQYDVMFLGTDIGTVLKVVSISKEKWNMEEVVLEELQIFKHSS  
IILNMELSLKQQQLYIGSRDGLVQLSLHRCDTYGKACADCCLARDPYCAWDGNACSRYPATSK  
RRARRQDVKYGDPITQCWDIEDSISHETADEKVI FGIEFNSTFLECI PKSQQATIKWYIQRSG  
DEHREELKPDERIIKTEYGLLIRSLQKKDSGMYCKAQEHTFIHTIVKLTLNVIENEQMENTQ  
RAEHEEGQVKDLLAESRLRYKDYIQILSSPNFSLDQYCEQMWHREKRRQRNKG GPKWKHMQEM  
KKKRNRHRDLDELPRAVAT

**Important features of the protein:****Signal peptide:**

amino acids 1-36

**N-glycosylation sites.**

amino acids 139-142, 607-610, 724-727

**Tyrosine kinase phosphorylation site.**

amino acids 571-576

**Gram-positive cocci surface proteins 'anchoring' hexapeptide.**

amino acids 32-37

352/615

**FIGURE 349**

CCCTGACCTCCCTGAGCCACACTGAGCTGGAAGCCGCAGAGGTCATCCTGGAGCATGCCCCACCGGGGAGCAGA  
CAACCTCCCAGGTAAGCTGGGAGCAAGACCTGAAGCTGTTTCTTCAGGAGCCTGGTGTATTTTCCCCACCCAC  
CTCAGCAGTTTCAGCCAGCAGGGACTGATCAGGTGTGTGTCTGGAGTGGGGAGCAGAAGGCGTGGCTGGCAAGA  
GTGGCCTGGAGAAAGAGGTTACGCGCTTGACCAGCCGAGCTGCCCGTGAACAAGATCCAGAACCATGGGCATC  
GGGTGAGGTGGGGGGGGCACAGGTGTCATGTGCACCTTCTTGTCTCAGCAAGAAGAGCTGAGAGAGGGGATCTTGG  
AGCCATTGAGGGTGTCTGAGCTACAGAGGGGAGGGAAGGTATTTTAAGGTAACAGTGTGGCACAATAGTTAA  
GAGCACAGTTTTTGGAGCTAGACCGACATAGGTTCAAATTCCTTCTGTTGCTTCTAGTTCTGTAGCCCCAGGT  
AAGGGAGTGACTTAACCTCTCTGGACTTCAATTTCTCATCTAAAGTAGGGCCAATAATAGCACCCACCTCAT  
AGGGAAGATTAAATGACATAATGTATGTGATGCACTAGCAAAAGTACCAGTCCCATAGTAAGTCATGCCCCACAG  
TATTTCCACCCACCCCTGTTCTCTGCCCTTCCCAACCAGGTAAGTGAACGACTGGAGCAGAGGGCGCAGCAGGCTT  
CAGAGCGGGAGGCTCCAAGCATAGAACAGAGGTTACAGGAAGTGCGAGAGAGCATCCGCCGGGCACAGGTGAGCC  
AGGTGAAGGGGGCTGCCCGGCTGGCCCTGCTGCAGGGGGCTGGCTTAGATGTGGAGCGTGGCTGAAGCCAGCCA  
TGACCCAGGCCCAGGATGAGGTGGAGCAGGAGCGGGCTCAGTGAGGCTCGGCTGTCCAGAGGGACCTCTCTC  
CAACCGCTGAGGATGCTGAGCTTTCTGACTTTGAGGAATGTGAGGAGACGGGAGAGCTCTTTGAGGAGCCTGCCC  
CCCAAGCCCTGGCCACGAGGGCCCTCCCTGCCCTGCACACGTGGTATTTCCGTATCAGGCAGGGCGTGAGGATG  
AGCTGACAATCACGGAGGGTGAAGTGGCTGGAGGTCATAGAGGAGGGAGATGCTGACGAATGGGTCAAGGCTCGGA  
ACCAGCACGGCGAGGTAGGCTTTGTCCCTGAGCGATATCTCAACTTCCCGGACCTCTCCCTCCCAGAGAGCAGCC  
AAGACAGTGACAATCCCTGCGGGGCAGAGCCACAGCATTCCTGGCACAGGCCCTGTACAGCTACACCGGACAGA  
GTGCAGAGGAGCTGAGCTTCCCTGAGGGGGCACTCATCCGTCTGCTGCCCGGGGCCAAGATGGAGTAGATGACG  
GCTTCTGGAGGGGAGAATTTGGGGGCCGTGTTGGGGTCTTCCCTCCCTGCTGGTGAAGAGCTGCTTGGCCCCC  
CAGGGCCACCTGAACCTCTCTGACCTGAACAGATGCTGCCGTCCCTTCTCCTCCAGCTTCTCCCCACCTGCAC  
CTACCTCTGTGTTGGATGGGCCCCCTGCACCTGTCTGCCCTGGGGACAAAGCCCTGGACTTCCCTGGGTTCTTG  
ACATGATGGCACCTCGACTCAGGCCGATGCGTCCACCACCTCCCCCGCCGGCTAAAGCCCCGGATCCCTGGCCACC  
CAGATCCCTCACCCTGAAGGCCAGGGAAGCCTTGACCCCAAGTGATGCTGCTGTCTTCAAGCTGTCAGA  
CCACACCATCAATGATCCAGAGCAACACAGCCAAAAGCTGGAATCGCCCTTATTTCCACCTCACCTCCAAGGCT  
GGAACTTGGCCCTTCCCATTTCTAGAGCTGGAACCCACTCCTTTTTTTCCCATTTGTTCTATCATCTCTAGGACC  
GGAATACTACCTTCTTCTGTGTGATGACCTATCTAGGCTGGTGAATGCCTGAAATCTCTGGGGCTGGAAACC  
ATCCATCAAGGTCTCTAGTAGTTCTGGCCACCTCTTCCCCACCTGGCTCCATGACCCACCCCACTCTGGATG  
CCAGGGTCACTGGGGTGGGCTGGGGAGAGGAACAGGCCCTGGGAATCAGGAGCTGGAGCCAGGATGCGAAGCAG  
CTGTAATGGTCTGAGCGGATTTATTGACAATGAATAAAGGGCACGAAGGCCAGGCCAGGGCTGGGCCCTTTGTG  
CTAAGAGGGCAGGGGGCTACGGTGTCTATTGCTTTAGGGGGCCACCAGGGCAGGGGCTGCTCCAGCTGCCAC  
GCTCTATCATATGGAGCGAGGTGTTGGGGAAGGCGGGGCAGGCAGCCTGTTGCAGGCAGGGGAAGGAGAAGAGAC  
TGAGGGGCTGTGACCTCTCTGAGGCCCCCAGCCTGAGACTGTGCAACTCCAGGTGGAAGTAGAGCTGGTCCCTC  
AGCTGGGGGGCAGTGCTGTCCAGTGGAGGGGAGGGCTTTACGCCCCACCCACCCCTGGCCCTGCCAGCTGGTAG  
TCCATCAGCACAAATGAAGGAGACTTGAGAGAAGAGGAAGAATAACACTGTTGCTTCTGTTCAAGCTGTGTCCAGC  
TTTTCCCTGGGGCTCCAGGACCTTCCCTACCTTGACCCACCAAGGATTTATAGCAAGGCTAAGCCTGC  
AGTTTACTCTGGGGGTTGAGGGAGCCGAAAGGCTTAAATAGTTTAAAGTAGGTGATGGGAAGATGAGATTACCTCA  
TTTAGGGCTCAGGCAGACTCACCTCACATACTCCCTGCTCCCTGTGGTAGAGACACCTGAGAGAAAGGGGAGGGG  
TCAACAATGAGAGACCAGGAGTAGGTCTATCAGTGCCCCCAGAGTAGAGAGCAATAAGAGCCCAGCCCAGTGC  
AGTCCCGCTGTGTTTTCTTACCTGGTGATCAGAAGTGTCTGGTTGCTTGGCTGCCCATTTGCCTCTTGAGTGG  
GCAGCCCTGGGCTTGGGCCCCCTCCCTCCGGCCCTCAGTGTTGGCTCTGCAGAAGCTCTGGGGTTCCCTCAAGTG  
CAGGAGGGGTTAGGCTGCTGTCCCTGAGTCTCCTATTCTGTACTGGGGGGCTGGCTAGGACCTGGGGCTGTGGCC  
TCTCAGGGGGCAGCCTCTCCATGGCAGGCATCCCTGCCCTGGGCTGCCCTCCCCCAGACCCCTGACCACCCCTG  
GGTCTGTCCCCCACCAGAGCCCCAGCTCCTGTCTGTGGGGGAGCCATCACGGTGTTCGTGCAGTCCATAGCGCT  
TCTCAATGTGTGTACCCGGAACCTGGGAGGGGAGGGAACACTGGGGTTTAGGACCACAACCTCAGAGGCTGCTTG  
GCCCTCCCCTCTGACCAGGGACATCCTGAGTTTGGTGGCTACTTCCCTCTGGCCTAAGGTAGGGGAGGCCCTTCTC  
AGATTGTGGGGCACATTGTGTAGCCTGACTTCTGCTGGAGCTCCAGTCCAGGAGGAAAGAGCCAAGGCCACTT  
TTGGGATCAGGTGCCTGATCACTGGGCCCCCTACCTCAGCCCCCTTCCCTGGAGACCTGCCCCACCTGCCCA  
CAGAGAACACAGTGTCTCCCTGTCCGGGGCGGCTTTTCTTCCCTTGGAGCGTCCCTGACGACCAAGTGGAG  
GCCTCTTGTGCGGCTGCAATGGATGCAAGGGGCTGCAGAGCCAGGTGCACTGTGTGATGATGGGAGGGGGCTC  
CGTCTGCAGGCTGGAGGTGGCATCCCACTGGACAGCAGGAGGAGGGGAGTGAGGGTAACATTTCCATTTCCCT  
TCATGTTTTGTTTTCTTACGTTCTTTCAGCATGCTCCTTAAACCCCAAGCCCCAATTTCCCAAGCCCCATTT  
TTTCTTGTCTTTATCTAATAAACTCAATATTAAG

353/615

**FIGURE 350**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73401
><subunit 1 of 1, 370 aa, 1 stop
><MW: 40685, pI: 4.53, NX(S/T): 0
MQLAKYQSHSKSCPTVFPPTPVLCLPNQVLQRLEQRRQQASEREAPSIEQRLQEVRESIRRAQ
VSQVKGAARLALLQGAGLDVERWLKPAMTQAQDEVEQERRLSEARLSQRDLSPATAEDAELSDF
EECEETGELFEPPAPQALATRALPCPAHVVFQYQAGREDELTITEGEWLEVIEEGDADEWVKA
RNQHGEVGFVPERYLNFPDLSLPESQDSNPGAEPTAFLAQALYSYTGQSAEELSFPEGAL
IRLLPRAQDGVDDGFWRGEFGGRVGVFPSLLVEELLGPPGPPPELSDPEQMLPSPSPPSFSPPA
PTSVLDGPPAPVLPQDKALDFPGFLDMMAPRLRPMRPPPPPPAKAPDPGHPDPLT
```

354/615

**FIGURE 351A**

CACAGGGAGACCCACAGACACATATGCACGAGAGAGACAGAGGAGGAAAGAGACAGAGACAAAGGCACAGCGGAA  
GAAGGCAGAGACAGGGCAGGCACAGAAGCGGCCAGACAGAGTCTACAGAGGGAGAGGCCAGAGAAGCTGCAGA  
AGACACAGGCAGGGAGAGACAAAGATCCAGGAAAGGAGGGCTCAGGAGGAGAGTTTGGAGAAGCCAGACCCCTGG  
GCACCTCTCCCAAGCCCAAGGACTAAGTTTTCTCCATTTCTTTAACGGTCTCAGCCCTTCTGAAAACCTTTGCC  
TCTGACCTTGGCAGGAGTCCAAGCCCCCAGGCTACAGAGAGAGCTTTCCAAAGCTAGGGTGTGGAGGACTTGGT  
GCCCTAGACGGCCTCAGTCCCTCCCAGCTGCAGTACCAGTGCCATGTCTCCAGACAGGCTCGCATCCCGGGAGGGG  
CTTGGCAGGGCGCTGGCTGTGGGGAGCCCAACCCTGCCTCCTGCTCCCCATTGTGCCGCTCTCCTGGCTGGTGTG  
GCTGCTTCTGCTACTGCTGGCCTCTCTCTGCCCTCAGCCCCGCTGGCCAGCCCCCTCCCCGGGAGGAGGAGAT  
CGTGTTCAGAGAAGCTCAACGGCAGCGTCTCTGCCCTCGGGCGCCCCCTGCCAGGCTGTTGTGCCGCTTGCA  
GGCCTTTGGGGAGACGCTGCTACTAGAGCTGGAGCAGGACTCCGGTGTGCAGGTCGAGGGGCTGACAGTGCAGTA  
CCTGGGCGAGGCGCCTGAGCTGCTGGGTGGAGCAGAGCCTGGCACCTACCTGACTGGCACCATCAATGGAGATCC  
GGAGTCGGTGGCATCTCTGCACTGGGATGGGGGAGCCCTGTAGCGCTGTTACAATATCGGGGGGCTGAAGTCCA  
CCTCCAGCCCCCTGGAGGGAGGCACCCCTAAGCTGCTGGGGGACCTGGGGCTCACATCTACGCCGGAAGAGTCC  
TGCCAGCGGTCAAGGTCCCATGTGCAACGTCAAGGCTCCTCTTGGAAGCCCCAGCCCCAGACCCGAAGAGCCAA  
GCGCTTTGCTTCACTGAGTAGATTTGTGGAGACACTGGTGGTGGCAGATGACAAGATGGCCGCATTCCACGGTGC  
GGGGCTAAAGCGCTACCTGCTAACAGTGATGGCAGCAGCAGCCAAGGCCTTCAAGCACCAAGCATCCGCAATCC  
TGTCAGCTTGGTGGTGACTCGGCTAGTGATCCTGGGGTCAGGCGAGGAGGGGCCCCAAGTGGGGCCAGTGTCTGC  
CCAGACCCCTGCGCAGCTTCTGTGCCTGGCAGCGGGGCCCTCAACACCCCTGAGGACTCGGGCCCTGACCACTTTGA  
CACAGCCATTCTGTTTACCCGTCAGGACCTGTGTGGAGTCTCCACTTGCACACGCTGGGTATGGCTGATGTGGG  
CACCGTCTGTGACCCGGCTCGGAGCTGTGCCATTGTGGAGGATGATGGGCTCCAGTCAGCCTTCACTGCTGCTCA  
TGAAGTGGGTGATGTCTTCAACATGCTCCATGACAACCTCAAGCCATGCATCAGTTTGAATGGGCCTTTGAGCAC  
CTCTCGCCATGTATGGCCCCCTGTGATGGCTCATGTGGATCCTGAGGAGCCCTGGTCCCCCTGCAGTGCCCGCTT  
CATCACTGACTTCTTGACAATGGCTATGGGCACTGTCTCTTAGACAAACCAGAGGCTCCATGCACTGCTGCTGT  
GACTTTCCCTGGCAAGGACTATGATGCTGACCGCTGACCTGACCTTCCGGGCCCTGACCCCTGCTGCTTCC  
ACAGCTGCCCGCCCTGTGCTGCCCTCTGGTGTCTGGCCACCTCAATGGCCATGCCATGTGCCAGACCAAAACA  
CTCGCCCTGGGCCGATGGCACACCCTGCGGGCCCGCACAGGCTGCATGGGTGGTCCGTGCCCTCCACATGGACCA  
GCTCCAGGACTTCAATATTCACAGGCTGGTGGCTGGGGTCCCTGGGGACCATGGGGTACTGCTCTCGGACCTG  
TGGGGGTGGTGTCCAGTTCTCTCCCAGACTGCACGAGGCCCTGTCCCCCGGAATGGTGGCAAGTACTGTGAGGG  
CCGCCGTACCCGCTTCCGCTCCTGCAACACTGAGGACTGCCCAACTGGCTCAGCCCTGACCTTCCGCGAGGAGCA  
GTGTGCTGCCTACAACCACCGCACCACCTCTTCAAGAGCTTCCAGGGCCCATGGACTGGGTTCCTCGCTACAC  
AGCGTGGCCCCCAGGACCTGCAAACTCACCTGCCAGGCCCGGGCACTGGGCTACTACTATGTGCTGGAGCC  
ACGGGTGGTAGATGGGACCCCTGTTCGCCGACAGCTCCTCGGTCTGTGTCCAGGGCCGATGCATCCATGCTGG  
CTGTGATCGCATATTGGCTCCAAGAAGAAGTTTGAAGAAGTGCATGGTGTGCGGAGGGGACGGTTCCTGGTTCAG  
CAAGCAGTCAGGCTCCTTCAGGAAATTACAGGTACGGATACAACAATGTGGTCACTATCCCCGCGGGGGCCACCCA  
CATTCTTGTCCGGCAGCAGGGAACCCCTGGCCACCCGAGCATCTACTTGGCCCTGAAGTGCAGATGGCTCCTA  
TGCCCTCAATGGTGAATACAGCTGATGCCCTCCCCCAGAGTGGTACTGCCTGGGGCAGTCAGCTTGCCTA  
CAGCGGGCCACTGCAGCCTCAGAGACACTGTAGGCCATGGGCCACTGGCCAGCCTTTGACACTGCAAGTCTCT  
AGTGGCTGGCAACCCCCAGGACACACGCCTCCGATACAGCTTCTTCGTGCCCCGGCCGACCCCTTCAACGCCACG  
CCCCACTCCCAGGACTGGCTGCACCGAAGAGCACAGATTCTGGAGATCCTTCGGCGGCGCCCCCTGGGCGGGCAG  
GAAATTAACCTCACTATCCCGCTGCCCTTTCTGGGCACCGGGGCTCGGACTTAGCTGGGAGAAAGAGAGAGCTT  
CTGTTGCTGCCTCATGCTAAGACTCAGTGGGGAGGGGCTGTGGGCGTGAGACCTGCCCCCTCTCTGCCCTAAT  
GCGCAGGCTGGCCCTGCCCTGGTTTCTGCCCTGGGAGGCAGTGATGGGTAGTGGATGGAGGGGCTGACAGAC  
AGCCCTCCATCTAAACTGCCCCCTCTGCCCTGCGGGTCACAGGAGGGAGGGGAAGGCAGGGAGGGCTGGGCCC  
CAGTTGTATTTATTTAGTATTTATTTACTTTTATTTAGCACAGGGAAGGGGACAAGGACTAGGGTCTGGGGAA  
CCTGACCCCTGACCCCTCATAGCCCTCACCTGGGGCTAGGAAATCCAGGGTGGTGGTATAGGTATAAGTGGTG  
TGTGTATGCGTGTGTGTGTGTGTGTGAAAATGTGTGTGTGCTTATGTATGAGGTACAACCTGTTCTGCTTTCTCT  
TTCCTGAATTTTATTTTGGGAAAAGAAAAGTCAAGGGTAGGGTGGGCCTTCAGGGAGTGAGGGATTATCTTTT  
TTTTTTTTTCTTTCTTTCTTTCTTTTGTGAGACAGAATCTCGCTCTGTGCGCCAGCTCGGAGTGAATG  
GCACAATCTCGGCTCACTGCATCCTCCGCCTCCCGGGTCAAGTGATTCTCATGCCTCAGCCTCTGAGTAGCTG  
GGATTACAGGCTCCTGCCACCACGCCCAGCTAATTTTGTGTTTGTGTTTGGAGACAGAGTCTCGCTATTGTC  
ACCAGGGCTGGAATGATTTAGCTCACTGCAACCTTCGCCACCTGGGTTCAGCAATTCTCTGCCCTCAGCCTCC  
CGAGTAGCTGAGATTATAGGCACCTACCACCACGCCCCGCTAATTTTGTATTTTGTATAGAGACGGGGTTTTCAC  
CATGTTGGCCAGGCTGGTCTCGAACTCCTGACCTTAGGTGATCCACTCGCCTTCATCTCCAAAGTGCTGGGATT  
ACAGGCGTGAGCCACCGTGCCTGGCCACGCCCCAATAATTTTGTATTTTGTATAGAGACAGGGTTTACCATGT  
TGGCCAGGCTGCTCTTGAACCTGACCTCAGGTAATCGACCTGCCTCGGCTCCCAAAGTGCTGGGATTACAGG

**FIGURE 351B**

[illegible]

356/615

**FIGURE 352**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73492

&gt;&lt;subunit 1 of 1, 837 aa, 1 stop

&gt;&lt;MW: 90167, pI: 8.39, NX(S/T): 1

MSQTGSHPGRGLAGRWLWGAQPCLLLPVPLSWLVWLLLLLLASLLPSARLASPLPREEEIVF  
PEKLNGSVLPGSGAPARLLCRLQAFGETLLLELEQDSGVQVEGLTVQYLGQAPELLGGAEPGT  
YLTGTINGDPESVASLHWDGGALLGVLYRGAEHLQLPLEGGTPNSAGGPGAHILRRKSPASG  
QGPMC NVKAPLGSPSPRRRAKRFASLSRFVETLVVADDKMAAFHGAGLKRYLLTVMAAAKA  
FKHPSIRNPVSLVVTRLVILGSGEEGPQVGPSAAQTLRSFCAWQRGLNTPEDSGPDHFDTAIL  
FTRQDL CGVSTCDTLGMADVGTVCDPARSCAIVEDDGLQSAFTAHELGHVFNMLHDNSKPCI  
SLNGPLSTSRHVMA PVMAHVDPEEPWSPCSARFITDFLDNGYGHCLLDKPEAPLHLPVTFPGK  
DYDADRQCQLTFGPDSRHCPQLPPPCAALWCSGHLNGHAMCQTKHSPWADGTPCGPAQACMGG  
RCLHMDQLQDFNIPQAGGWGPWGPWGDCSRTC GGGVQFSSRDCTRVPVRNGGKYCEGRTRFR  
SCNTEDCPTGSALTFREEQCAAYNHRTDLFKSFPGPMDWVPRYTGVAPQDQCKLTCQARALGY  
YYVLEPRVVDGTPCSPDSSSV CVQGRCIHAGCDRIIGSKKKFDKCMVCGGDGSGCSKQSGSFR  
KFRYGYNNVVTIPAGATHILVRQQGNPGHRSIYLALKLPDGSYALNGEYTLMPSPTDVVLPGA  
VSLRYSGATAASETLSGHGPLAQPLTLQVLVAGNPQDTRLRYSFFVPRPTPSTPRPTPDWLH  
RRAQILEILRRRPWAGRK

**Important features of the protein:****Signal peptide:**

amino acids 1-48

**N-glycosylation site.**

amino acids 68-71

**Glycosaminoglycan attachment site**

amino acids 188-191, 772-775

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 182-185

**Tyrosine kinase phosphorylation site.**

amino acids 730-736

**N-myristoylation sites.**

amino acids 5-10, 19-24, 121-126, 125-130, 130-135, 147-152, 167-172, 168-173, 174-179, 323-328, 352-357, 539-544, 555-560, 577-582, 679-684, 682-687, 763-768

**Amidation sites.**

amino acids 560-563, 834-837

**Leucine zipper pattern.**

amino acids 17-38, 24-45

**Neutral zinc metallopeptidases, zinc-binding region signature.**

amino acids 358-367

357/615

**FIGURE 353**

GCGGAACTGGCTCCGGCTGGCACCTGAGGAGCGGCGTGACCCCGAGGGCCCAGGGAGCTGCCC  
GGCTGGCCTAGGCAGGCAGCCGCACCATGGCCAGCACGGCCGTGCAGCTTCTGGGCTTCCTGC  
TCAGCTTCCTGGGCATGGTGGGCACGTTGATCACCACCATCCTGCCGCACTGGCGGAGGACAG  
CGCACGTGGGCACCAACATCCTCACGGCCGTGTCCTACCTGAAAGGGCTCTGGATGGAGTGTG  
TGTGGCACAGCACAGGCATCTACCAGTGCCAGATCTACCGATCCCTGCTGGCGCTGCCCCAAG  
ACCTCCAGGCTGCCCCGCGCCCTCATGGTCATCTCCTGCCTGCTCTCGGGCATAGCCTGCGCCT  
GCGCCGTCATCGGGATGAAGTGCACGCGCTGCGCCAAGGGCACACCCGCCAAGACCACCTTTG  
CCATCCTCGGCGGCACCCTCTTCATCCTGGCCGGCCTCCTGTGCATGGTGGCCGTCTCCTGGA  
CCACCAACGACGTGGTGCAGAACTTCTACAACCCGCTGCTGCCAGCGGCATGAAGTTTGAGA  
TTGGCCAGGCCCTGTACCTGGGCTTCATCTCCTCGTCCCTCTCGCTCATTGGTGGCACCCCTGC  
TTTGCTGTCTGCCAGGACGAGGCACCCTACAGGCCCTACCAGGCCCCGCCAGGGCCACCA  
CGACCACTGCAAACACCGCACCTGCCTACCAGCCACCAGCTGCCTACAAAGACAATCGGGCCC  
CCTCAGTGACCTCGGCCACGCACAGCGGGTACAGGCTGAACGACTACGTGTGAGTCCCCACAG  
CCTGCTTCTCCCCTGGGCTGCTGTGGGCTGGGTCCCCGGCGGGACTGTCAATGGAGGCAGGGG  
TTCCAGCACAAAGTTTACTTCTGGGCAATTTTTGTATCCAAGGAAATAATGTGAATGCGAGGA  
AATGTCTTTAGAGCACAGGGACAGAGGGGGAAATAAGAGGAGGAGAAAGCTCTCTATACCAA  
GACTGAAAAAAAAAATCCTGTCTGTTTTTGTATTTATTATATATATTTATGTGGGTGATTTGA  
TAACAAGTTTAATATAAAGTGACTTGGGAGTTTGGTCAGTGGGGTTGGTTTGTGATCCAGGAA  
TAAACCTGCGGATGTGGCTGTTTATGAAAAAAAAAAAAA



358/615

## **FIGURE 354**

MASTAVQLLGFLLSFLGMVGTLLITTLPHWRRTAHVGTNILTAVSYLKGLWMECVWHSTGIYQ  
CQIYRSLALPQDLQAARALMVISCLLSGIACACAVIGMKCTRCAKGTPAKTTFAILGGTLFI  
LAGLLCMVAVSWTTNDVVQNFYNPLPSGMKFEIGQALYLGFISSLSLIGGTLLCLSCQDEA  
PYRPYQAPPRATTTTANTAPAYQPPAAYKDNRAPSVTSATHSGYRLNDYV

**Important features of the protein:**

**Signal peptide:**

amino acids 1-21

**Transmembrane domains:**

amino acids 82-103, 115-141, 160-182

359/615

**FIGURE 355**

GAGCTCCCCCTCAGGAGCGCGTTAGCTTCACACCTTCGGCAGCAGGAGGGCGGCAGCTTCTCGC  
AGGCGGCAGGGCGGGCGGCCAGGATC**ATG**TCCACCACCACATGCCAAGTGGTGGCGTTCCCTCC  
TGTCCATCCTGGGGCTGGCCGGCTGCATCGCGGCCACCGGGATGGACATGTGGAGCACCCAGG  
ACCTGTACGACAACCCCGTCACCTCCGTGTTCCAGTACGAAGGGCTCTGGAGGAGCTGCGTGA  
GGCAGAGTTTCAGGCTTCACCGAATGCAGGCCCTATTTACCACATCCTGGGACTTCCAGCCATGC  
TGCAGGCAGTGCAGGCCCTGATGATCGTAGGCATCGTCCTGGGTGCCATTGGCCTCCTGGTAT  
CCATCTTTGCCCTGAAATGCATCCGCATTGGCAGCATGGAGGACTCTGCCAAAGCCAACATGA  
CACTGACCTCCGGGATCATGTTCAATTGTCTCAGGTCTTTGTGCAATTGCTGGAGTGTCTGTGT  
TTGCCAACATGCTGGTGACTAACTTCTGGATGTCCACAGCTAACATGTACACCGGCATGGGTG  
GGATGGTGCAGACTGTTTCAGACCAGGTACACATTTGGTGGCGCTCTGTTTCGTGGGCTGGGTG  
CTGGAGGCCTCACACTAATTGGGGGTGTGATGATGTGCATCGCCTGCCGGGGCCTGGCACCAG  
AAGAAACCAACTACAAAGCCGTTTCTTATCATGCCTCAGGCCACAGTGTTGCCTACAAGCCTG  
GAGGCTTCAAGGCCAGCACTGGCTTTGGGTCCAACACCAAAAACAAGAAGATATACGATGGAG  
GTGCCCCGACAGAGGACGAGGTACAATCTTATCCTTCCAAGCACGACTATGTG**TAA**TGCTCTA  
AGACCTCTCAGCACGGGCGGAAGAACTCCCGGAGAGCTCACCCAAAAACAAGGAGATCCCA  
TCTAGATTTCTTCTTGCTTTTGACTCACAGCTGGAAGTTAGAAAAGCCTCGATTTTCATCTTTG  
GAGAGGCCAAATGGTCTTAGCCTCAGTCTCTGTCTCTAAATATTCCACCATAAAACAGCTGAG  
TTATTTATGAATTAGAGGCTATAGCTCACATTTTCAATCCTCTATTTCTTTTTTAAATATAA  
CTTTCTACTCTGATGAGAGAATGTGGTTTTAATCTCTCTCTCACATTTTGATGATTTAGACAG  
ACTCCCCCTCTTCCTCCTAGTCAATAAACCCATTGATGATCTATTTCCAGCTTATCCCCAAG  
AAAACCTTTTGAAGGAAAGAGTAGACCCAAAGATGTTATTTTCTGCTGTTTGAATTTTGTCTC  
CCCACCCCCAACTTGGCTAGTAATAAACACTTACTGAAGAAGAAGCAATAAGAGAAAGATATT  
TGTAATCTCTCCAGCCCATGATCTCGGTTTTCTTACACTGTGATCTTAAAAGTTACCAAACCA  
AAGTCATTTTCAGTTTGAGGCAACCAAACCTTTCTACTGCTGTTGACATCTTCTTATTACAGC  
AACACCATTCTAGGAGTTTCTGAGCTCTCCACTGGAGTCCTCTTTCTGTCGCGGGTCAGAAA  
TTGTCCCTAGATGAATGAGAAAATTATTTTTTTTAAATTTAAGTCCTAAATATAGTTAAAATAA  
ATAATGTTTTAGTAAAATGATACACTATCTCTGTGAAATAGCCTCACCCCTACATGTGGATAG  
AAGGAAATGAAAAAATAATTGCTTTGACATTGTCTATATGGTACTTTGTAAAGTCATGCTTAA  
GTACAAATTCATGAAAAGCTCACACCTGTAATCCTAGCACTTTGGGAGGCTGAGGAGGAAGG  
ATCACTTGAGCCCAGAAGTTCGAGACTAGCCTGGGCAACATGGAGAAGCCCTGTCTCTACAAA  
ATACAGAGAGAAAAAATCAGCCAGTCATGGTGGCATAACCTGTAGTCCCAGCATTCCGGGAG  
GCTGAGGTGGGAGGATCACTTGAGCCCAGGGAGGTTGGGGCTGCAGTGAGCCATGATCACACC  
ACTGCACTCCAGCCAGGTGACATAGCGAGATCCTGTCTAAAAAAATAAAAAATAAATAATGGA  
ACACAGCAAGTCCTAGGAAGTAGGTTAAAACCTAATTCTTTAA

360/615

**FIGURE 356**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73734

&gt;&lt;subunit 1 of 1, 261 aa, 1 stop

&gt;&lt;MW: 27856, pI: 8.50, NX(S/T): 1

MSTTTCQVVAFLLSILGLAGCIAATGMDMWSTQDLYDNPVTSVVFQYEGLRSCVRQSSGFTEC  
RPYFTILGLPAMLQAVRALMIVGIVLGAIGLLVSIFALKCIRIGSMEDSAKANMTLTSGIMFI  
VSGLCAIAGVSVFANMLVTNFWMSTANMYTGMGGMVQTVQTRYTFGAALFVGWVAGGLTLIGG  
VMMCIACRGLAPEETNYKAVSYHASGHSVAYKPGGFKASTGFGSNTKNKKIYDGGARTEDEVQ  
SYPSKHDYV

**Signal peptide:**

amino acids 1-23

**Transmembrane domains:**

amino acids 81-100, 121-141, 173-194

361/615

**FIGURE 357**

GGAAAACTGTTCTCTTCTGTGGCACAGAGAACCCTGCTTCAAAGCAGAAGTAGCAGTTCCGG  
AGTCCAGCTGGCTAAAACTCATCCCAGAGGATA**ATG**GCAACCCATGCCTTAGAAATCGCTGGG  
CTGTTTCTTGGTGGTGTGGAATGGTGGGCACAGTGGCTGTCACTGTCATGCCTCAGTGGAGA  
GTGTCGGCCTTCATTGAAAACAACATCGTGGTTTTTGAAGAACTTCTGGGAAGGACTGTGGATG  
AATTGCGTGAGGCAGGCTAACATCAGGATGCAGTGCAGAAATCTATGATTCCCTGCTGGCTCTT  
TCTCCGGACCTACAGGCAGCCAGAGGACTGATGTGTGCTGCTTCCGTGATGTCCTTCTTGGCT  
TTCATGATGGCCATCCTTGGCATGAAATGCACCAGGTGCACGGGGGACAATGAGAAGGTGAAG  
GCTCACATTCTGCTGACGGCTGGAATCATCTTCATCATCACGGGCATGGTGGTGTCTATCCCT  
GTGAGCTGGGTTGCCAATGCCATCATCAGAGATTTCTATAACTCAATAGTGAATGTTGCCCAA  
AAACGTGAGCTTGGAGAAGCTCTCTACTTAGGATGGACCACGGCACTGGTGTGATTGTTGGA  
GGAGCTCTGTTCTGCTGCGTTTTTTGTGCAACGAAAAGAGCAGTAGCTACAGATACTCGATA  
CCTTCCCATCGCACAAACCCAAAAAAGTTATCACACCGGAAAGAAGTCACCGAGCGTCTACTCC  
AGAAGTCAGTATGTG**TAG**TTGTGTATGTTTTTTAACTTTACTATAAAGCCATGCAATGACA  
AAAATCTATATTACTTTCTCAAAATGGACCCCAAAGAACTTTGATTTACTGTTCTTAACTGC  
CTAATCTTAATTACAGGAAGTGTGCATCAGCTATTTATGATTCTATAAGCTATTTTCAGCAGAA  
TGAGATATTAAACCCAATGCTTTGATTGTTCTAGAAAGTATAGTAATTTGTTTTCTAAGGTGG  
TTCAAGCATCTACTCTTTTTATCATTACTTTCAAAATGACATTGCTAAAGACTGCATTATTTT  
ACTACTGTAATTTCTCCACGACATAGCATTATGTACATAGATGAGTGTAAACATTTATATCTCA  
CATAGAGACATGCTTATATGGTTTTATTTAAAATGAAATGCCAGTCCATTACACTGAATAAAT  
AGAAGTCAACTATTGCTTTTCAGGGAAATCATGGATAGGGTTGAAGAAGGTACTATTAATTG  
TTTAAAAACAGCTTAGGGATTAATGTCCTCCATTTATAATGAAGATTAAAATGAAGGCTTTAA  
TCAGCATTGTAAAGGAAATTGAATGGCTTTCTGATATGCTGTTTTTTAGCCTAGGAGTTAGAA  
ATCCTAACTTCTTTATCCTCTTCTCCCAGAGGCTTTTTTTTTCTTGTGTATTAAATTAACATT  
TTTAAAACGCAGATATTTTGTCAAGGGGCTTTGCATTCAAAGTCTTTTCCAGGGCTATACTC  
AGAAGAAAGATAAAAAGTGTGATCTAAGAAAAAGTGTGGTTTTTAGGAAAGTGAAAATATTTTT  
GTTTTTGTATTTGAAGAAGAATGATGCATTTTGACAAGAAATCATATATGTATGGATATATTT  
TAATAAGTATTTGAGTACAGACTTTGAGGTTTCATCAATATAAATAAAAAGAGCAGAAAAATAT  
GTCTTGGTTTTTCATTTGCTTACCAAAAAACAACAACAAAAAAGTTGTCCTTTGAGAACTTC  
ACCTGCTCCTATGTGGGTACCTGAGTCAAAATTGTCATTTTTGTTCTGTGAAAAATAAATTTT  
CTTCTTGTAACATTTCTGTTTAGTTTTACTAAAATCTGTAAATACTGTATTTTTCTGTTTATT  
CCAAATTTGATGAACTGACAATCCAATTTGAAAGTTTGTGTGACGTCTGTCTAGCTTAAAT  
GAATGTGTTCTATTTGCTTTATACATTTATATTAATAAATTGTACATTTTTTCTAATT

362/615

**FIGURE 358**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73735

&gt;&lt;subunit 1 of 1, 225 aa, 1 stop

&gt;&lt;MW: 24845, pI: 9.07, NX(S/T): 0

MATHALEIAGLFLGGVGMVGTVAVTVMPOWRVSFAFIENNIVVFENFWEGLWMNCVRQANIRMQ  
CKIYDSLLALSPDLQAARGLMCAASVMSFLAFMMAILGMKCTRCTGDNEKVKAHILLTAGIIF  
IITGMVVLIPVSWVANAIIRDFYNSIVNVAQKRELGEALYLGWTTALVLIVGGALFCCVFCCN  
EKSSSYRYSIPSHRTTQKSYHTGKKSPSVYSRSQYV

**Signal peptide:**

amino acids 1-17

**Transmembrane domains:**

amino acids 82-101, 118-145, 164-188

363/615

**FIGURE 359**

CCCGCGCCCGGTTCTCCCTCGCAGCACCTCGAAGTGCGCCCCCTCGCCCTCCTGCTCGCGCCCC  
GCCGCC**ATG**GCTGCCTCCCCGCGCGGCCTGCTGTCCTGGCCCTGACCGGGCTGGCGCTGCTC  
CTGCTCCTGTGCTGGGGCCCAGGTGGCATAAGTGGAATAAACTCAAGCTGATGCTTCAAAAA  
CGAGAAGCACCTGTTCCAATAAGACTAAAGTGGCCGTTGATGAGAATAAAGCCAAAGAATTC  
CTTGGCAGCCTGAAGCGCCAGAAGCGGCAGCTGTGGGACCGGACTCGGCCCGAGGTGCAGCAG  
TGGTACCAGCAGTTTCTCTACATGGGCTTTGATGAAGCGAAATTTGAAGATGACATCACCTAT  
TGGCTTAACAGAGATCGAAATGGACATGAATACTATGGCGATTACTACCAACGTCACTATGAT  
GAAGACTCTGCAATTGGTCCCCGAGCCCCCTACGGCTTTAGGCATGGAGCCAGCGTCAACTAC  
GATGACTAC**TAA**CCATGACTTGCCACACGCTGTACAAGAAGCAAATAGCGATTCTCTTCATGT  
ATCTCCTAATGCCTTACACTACTTGGTTTCTGATTTGCTCTATTTTCAGCAGATCTTTTCTACC  
TACTTTGTGTGATCAAAAAAGAAGAGTTAAAACAACACATGTAAATGCCTTTTGATATTTTCAT  
GGGAATGCCTCTCATTTAAAAATAGAAATAAAGCATTTTGTTAAAAAGA

364/615

**FIGURE 360**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73742

&gt;&lt;subunit 1 of 1, 148 aa, 1 stop

&gt;&lt;MW: 17183, pI: 8.77, NX(S/T): 0

MAASPARPAVLALTGLALLLLLCWGPGGISGNKCLKMLQKREAPVPTKTKVAVDENKAKEFLG

SLKRQKRQLWDRTRPEVQQWYQQFLYMGFDEAKFEDDITYWLNDRDRNGHEYYGDIYQRHYDED

SAIGPRSPYGFRHGASVNYDDY

**Signal peptide:**

amino acids 1-30





366/615

**FIGURE 362**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73746

&gt;&lt;subunit 1 of 1, 148 aa, 1 stop

&gt;&lt;MW: 16896, pI: 6.05, NX(S/T): 1

MTKALLIYLVSSFLALNQASLISRCDLAQVLQLEDLDGFEGLSDWLCIAFVESKFNISKIN

ENADGSFDYGLFQINSHYWCNDYKSYSENLCVDCQDLLNPNNLAGIHCAKRIVSGARGMNNW

VEWRLHCSGRPLSYWLTGCRLR

**Signal peptide:**

amino acids 1-18

367/615

**FIGURE 363**

TCTGACCTGACTGGAAGCGTCCAAAGAGGGACGGCTGTCAGCCCTGCTTGACTGAGAACCCAC  
CAGCTCATCCCAGACACCTCATAGCAACCTATTTATACAAAGGGGGAAAGAAACACCTGAGCA  
GAATGGAATCATTATTTTTTTTCCCAAGGAGAAAACCGGGGTAAAGGGAGGGAAGCAATTCAAT  
TTGAAGTCCCTGTGAATGGGCTTTCAGAAGGCAATTAAAGAAATCCACTCAGAGAGGACTTGG  
GGTGAAACTTGGGTCTGTGGTTTTCTGATTGTAAGTGGAAGCAGGTCTTGACACGCTGTTG  
GCAAATGTCAGGACCAGGTAAAGTGAAGTGGCAGAAAACTTCCAGGTGGAACAAGCAACCCAT  
GTTCTGCTGCAAGCTTGAAGGAGCCTGGAGCGGGAGAAAGCTAACTTGAACATGACCTGTTGC  
ATTTGGCAAGTTCTAGCAACATGCTCCTAAGGAAGCGATACAGGCACAGACCATGCAGACTCC  
AGTTCCTCCTGCTGCTCCTGATGCTGGGATGCGTCTGATGATGGTGGCGATGTTGCACCCTC  
CCCACCACACCCTGCACCAGACTGTACAGCCCAAGCCAGCAAGCACAGCCCTGAAGCCAGGT  
ACCGCCTGGACTTTGGGGAATCCCAGGATTGGGTACTGGAAGCTGAGGATGAGGGTGAAGAGT  
ACAGCCCTCTGGAGGGCCTGCCACCCTTTATCTCACTGCGGGAGGATCAGCTGCTGGTGGCCG  
TGGCCTTACCCAGGCCAGAAGGAACAGAGCCAGGGCAGGAGAGGTGGGAGCTACCGCCTCA  
TCAAGCAGCCAAGGAGGCAGGATAAGGAAGCCCCAAAGAGGGACTGGGGGGCTGATGAGGACG  
GGGAGGTGTCTGAAGAAGAGGAGTTGACCCCGTTTACGCTGGACCCACGTGGCCTCCAGGAGG  
CACTCAGTGGCCGCATCCCCCTCCAGAGGGCTCTGCCCAGGTGCGGCACCCACTGTGTCTGC  
AGCAGCACCCCTCAGGACAGCCTGCCACAGCCAGCGTCATCCTCTGTTTCCATGATGAGGCCT  
GGTCCACTCTCCTGCGGACTGTACACAGCATCCTCGACACAGTGGCCAGGGCCTTCTGAAGG  
AGATCATCCTCGTGGACGACCTCAGCCAGCAAGGACAACCTCAAGTCTGCTCTCAGCGAATATG  
TGGCCAGGCTGGAGGGGGTGAAGTTACTCAGGAGCAACAAGAGGCTGGGTGCCATCAGGGCCC  
GGATGCTGGGGGCCACCAGAGCCACCGGGGATGTGCTCGTCTTCATGGATGCCCACTGCGAGT  
GCCACCCAGGCTGGCTGGAGCCCCTCCTCAGCAGAATAGCTGGTGACAGGAGCCGAGTGGTAT  
CTCCGGTGATAGATGTGATTGACTGGAAGACTTTCCAGTATTACCCCTCAAAGGACCTGCAGC  
GTGGGGTGTGGACTGGAAGCTGGATTTCCACTGGGAACCTTTGCCAGAGCATGTGAGGAAGG  
CCCTCCAGTCCCCCATAAGCCCCATCAGGAGCCCTGTGGTGCCCGGAGAGGTGGTGGCCATGG  
ACAGACATTACTTCCAAAACACTGGAGCGTATGACTCTCTTATGTGCTGCGAGGTGGTGAAA  
ACCTCGAACTGTCTTTCAAGGCCTGGCTCTGTGGTGGCTCTGTTGAAATCCTTCCCTGCTCTC  
GGGTAGGACACATCTACCAAATCAGGATTCCATTCCCCCTCGACCAGGAGGCCACCCTGA  
GGAACAGGGTTTCGCTTGCTGAGACCTGGCTGGGGTCATTCAAAGAAACCTTCTACAAGCATA  
GCCAGAGGCCTTCTCCTTGAGCAAGGCTGAGAAGCCAGACTGCATGGAACGCTTGACGCTGC  
AAAGGAGACTGGGTTGTGCGACATTCCACTGGTTTCTGGCTAATGTCTACCCTGAGCTGTACC  
CATCTGAACCCAGGCCAGTTTCTCTGGAAAGCTCCACAACACTGGACTTGGGCTCTGTGCAG  
ACTGCCAGGCAGAAGGGGACATCCTGGGCTGTCCATGGTGTGGCTCCTTGCAAGTGCAGGCC  
GGCAGCAACAGTACCTGCAGCACACCAGCAGGAAGGAGATTCACTTTGGCAGCCACAGCACC  
TGTGCTTTGCTGTCAGGCAGGAGCAGGTGATTCTTCAGAACTGCACGGAGGAAGGCCTGGCCA  
TCCACCAGCAGCACTGGGACTTCCAGGAGAATGGGATGATTGTCCACATTCTTTCTGGGAAAT  
GCATGGAAGCTGTGGTGCAAGAAAACAATAAAGATTTGTACCTGCGTCCGTGTGATGGAAAAG  
CCCGCCAGCAGTGGCGATTTGACCAGATAAATGCTGTGGATGAACGATGCAATGTCAATGTCAG  
AAGGAAAAGAGAATTTTGGCCATCAAAATCCAGCTCCAAGTGAACGTAAAGAGCTTATATATT  
TCATGAAGCTGATCCTTTTGTGTGTGTGCTCCTTGTGTTAGGAGAGAAAAAGCTCTATGAAA  
GAATATAGGAAGTTTCTCCTTTTACACCTATTTTCATTGACTGCTGGCTGCTTA

368/615

**FIGURE 364**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA73760

&gt;&lt;subunit 1 of 1, 639 aa, 1 stop

&gt;&lt;MW: 73063, pI: 6.84, NX(S/T): 2

MLLRKRYRHRPCRLQFLLLLLMLGCVLMMVAMLHPPHHTLHQTVTAQASKHSPEARYRLDFGE  
SQDWVLEAEDEGE EYSPL EGLPPFISLREDQLLVAVALPQARRNQSQGRRGGSYRLIKQPRRQ  
DKEAPKRDWGADEDEGEVSEEEELTPFSLDPRGLQEALSARIPLQRALPEVRHPLCLQQHPQDS  
LPTASVILCFHDEAWSTLLRTVHSILDTVPRAFLKEIILVDDL SQQQLKSALSEYVARLEGV  
KLLRSNKRLGAIRARMLGATRATGDVLVFMDAHCECHPGWLEPLLSRIAGDRSRVVSPVIDVI  
DWKTFQYYPSKDLQRGVLDWKLDHFHWEPLPEHVRKALQSPISPIRSPVVPGEVVAMDRHYFQN  
TGAYDSLMSLRGGENLELSFKAWLCGGSVEILPCSRVGHYIYNQDSHSPLDQEATLRNRVRIA  
ETWLGSFKETFYKHSPEAFSLSKAEKPCMERLQLQRRLGCRTFHWFLANVYPELYPSEPRPS  
FSGKLHNTGLGLCADCQAEGDILGCPMVLAPCSDSRQQQYLQHTSRKEIHFGSPQHLCFAVRQ  
EQVILQNCTEEGLAIHQHWFQENGMIVHILSGKCMEAVVQENNKDLYLRPCDGKARQQWRF  
DQINAVDER

**Signal peptide:**

amino acids 1-28

369/615

**FIGURE 365**

GGAGAGAGGCGCGCGGGTGAAAGGCGCATTGATGCAGCCTGCGGCGGCCTCGGAGCGCGGCGG  
AGCCAGACGCTGACCACGTTCTCTCCTCGGTCTCCTCCGCCTCCAGCTCCGCGCTGCCCCGGC  
AGCCGGGAGCC**ATG**CGACCCAGGGCCCCGCGCCTCCCCGCAGCGGCTCCGCGGCCTCCTGC  
TGCTCCTGCTGCTGCAGCTGCCCCGCGCCGTCGAGCGCCTCTGAGATCCCCAAGGGGAAGCAAA  
AGGCGCAGCTCCGGCAGAGGGAGGTGGTGGACCTGTATAATGGAATGTGCTTACAAGGGCCAG  
CAGGAGTGCTTGGTCGAGACGGGAGCCCTGGGGCCAATGTTATTCCGGGTACACCTGGGATCC  
CAGGTCGGGATGGATTCAAAGGAGAAAAGGGGGAATGTCTGAGGGAAAGCTTTGAGGAGTCCT  
GGACACCCAACTACAAGCAGTGTTTCATGGAGTTCATTGAATTATGGCATAGATCTTGGGAAAA  
TTGCGGAGTGACATTTACAAAGATGCGTTCAAATAGTGCTCTAAGAGTTTTGTTCAGTGGCT  
CACTTCGGCTAAAATGCAGAAATGCATGCTGTCAGCGTTGGTATTTACATTCAATGGAGCTG  
AATGTTTCAGGACCTCTTCCCATTGAAGCTATAATTTATTTGGACCAAGGAAGCCCTGAAATGA  
ATTCAACAATTAATATTCATCGCACTTCTTCTGTGGAAGGACTTTGTGAAGGAATTGGTGCTG  
GATTAGTGGATGTTGCTATCTGGGTTGGCACTTGTTTCAGATTACCCAAAAGGAGATGCTTCTA  
CTGGATGGAATTCAGTTTCTCGCATCATTATTGAAGAACTACCAAAA**TAA**ATGCTTTAATTTT  
CATTTGCTACCTCTTTTTTTTATTATGCCTTGGAATGGTTCACTTAAATGACATTTTAAATAAG  
TTTATGTATACATCTGAATGAAAAGCAAAGCTAAATATGTTTACAGACCAAAGTGTGATTTCA  
CACTGTTTTTTAAATCTAGCATTATTCATTTTGCTTCAATCAAAGTGGTTTCAATATTTTTTT  
TAGTTGGTTAGAATACTTTCTTCATAGTCACATTCTCTCAACCTATAATTTGGAATATTGTTG  
TGGTCTTTTGTTTTTTCTCTTAGTATAGCATTTTTAAAAAATATAAAAGCTACCAATCTTTG  
TACAATTTGTAAATGTTAAGAATTTTTTTTATATCTGTAAATAAAAATTATTTCCAACA

370/615

**FIGURE 366**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76393

&gt;&lt;subunit 1 of 1, 243 aa, 1 stop

&gt;&lt;MW: 26266, pI: 8.43, NX(S/T): 1

MRPQGPAASPQRLRGLLLLLLLQLPAPSSASEIPKGKQKAQLRQREVVDLYNGMCLQGPAGVP  
GRDGSPGANVIPGTPGIPGRDGFKEKGECLRESFEESWTPNYKQCSWSSLNYGIDLGKIAEC  
TFTKMRSNSALRVLFSGSLRLKCRNACCQRWYFTFNGAECGPLPIEAIYLDQGSPPEMNSTI  
NIHRTSSVEGLCEGIGAGLVDVAIWVGTCSDYPKGDASTGWNSVSRIIIEELPK

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 195-217

371/615

**FIGURE 367**

GTAAACCAGCGCAGTCCTCCGTGCGTCCCGCCCGCCGCTGCCCTCACTCCCGGCCAGG**ATGGC**  
ATCCTGTCTGGCCCTGCGCATGGCGCTGCTGCTGGTCTCCGGGGTTCTGGCCCCTGCGGTGCT  
CACAGACGATGTTCCACAGGAGCCCGTGCCACGCTGTGGAACGAGCCGGCCGAGCTGCCGTC  
GGGAGAAGGCCCCGTGGAGAGCACCCAGCCCCGGCCGGGAGCCCGTGGACACCGGTCCCCCAGC  
CCCCACCGTCGCGCCAGGACCCGAGGACAGCACCGCGCAGGAGCGGCTGGACCAGGGCGGCGG  
GTCGCTGGGGCCCGGCGCTATCGCGGCCATCGTGATCGCCGCCCTGCTGGCCACCTGCGTGGT  
GCTGGCGCTCGTGGTCGTCGCGCTGAGAAAGTTTTCTGCCTCC**TGA**AGCGAATAAAGGGGCCG  
CGCCCGCCGCGGCGCGACTCGGCAAAAAAAAAAAAAAA

372/615

**FIGURE 368**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76398

&gt;&lt;subunit 1 of 1, 121 aa, 1 stop

&gt;&lt;MW: 12073, pI: 4.11, NX(S/T): 0

MASCLALRMALLLVSGVLAPAVLTDDVPQEPVPTLWNEPAELPSGEGPVVESTSPGREPVDTGP

PAPTVPAGPEDSTAQERLDQGGGSLGPGAIAAIVIAALLATCVVLALVVVALRKFSAS

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 91-110

**Glycosaminoglycan attachment site.**

amino acids 44-47

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 116-119

**N-myristoylation site.**

amino acids 91-96

373/615

**FIGURE 369**

GGCCGTTGGTTGGTGCGCGGCTGAAGGGTGTGGCGCGAGCAGCGTCGTTGGTTGGCCGGCGGC  
GGGCCGGGACGGGC**ATG**GCCCTGCTGCTGTGCCTGGTGTGCCTGACGGCGGCGCTGGCCCACG  
GCTGTCTGCACTGCCACAGCAACTTCTCCAAGAAGTTCTCCTTCTACCGCCACCATGTGAACT  
TCAAGTCTTGGTGGGTGGGCGACATCCCCGTGTCAGGGGCGCTGCTCACCGACTGGAGCGACG  
ACACGATGAAGGAGCTGCACCTGGCCATCCCCGCCAAGATCACCCGGGAGAAGCTGGACCAAG  
TGGCGACAGCAGTGTACCAGATGATGGATCAGCTGTACCAGGGGAAGATGTACTTCCCCGGGT  
ATTTCCCCAACGAGCTGCGAAACATCTTCCGGGAGCAGGTGCACCTCATCCAGAACGCCATCA  
TCGAAAGGCACCTGGCACCAGGCAGCTGGGGAGGAGGGCAGCTCTCCAGGGAGGGACCCAGCC  
TAGCACCTGAAGGATCAATGCCATCACCCCGCGGGGACCTCCCC**TAA**GTAGCCCCCAGAGGGCG  
CTGGGAGTGTGGCACCGCCCTCCCCTGAAGTTTGCTCCATCTCACGCTGGGGGTCAACCTGG  
GGACCCCTTCCCTCCGGGCCATGGACACACATACATGAAAACCAGGCCGCATCGACTGTCAGC  
ACCGCTGTGGCATCTTCCAGTACGAGACCATCTCCTGCAACAACCTGCACAGACTCGCACGTGC  
CCTGCTTTGGCTATAACTGCGAGTAGGGCTCAGGCATCACACCCACCCGTGCCAGGGCCCTAC  
TGTCCCTGGGGTCCCAGGCTCTCCTTGGAGGGGGCTCCCCGCCTTCCACCTGGCTGTATCGG  
GTAGGGCGGGGCCGTGGGTTCAGGGGCGCACCACTTCCAAGCCTGTGTCCACAGGTCTTCGG  
CGCAGTGGAAGTCAGCTGTCCAGGGCCTCCTGAACTACATAAATAACTGGCACAAGTAAGTCC  
CCTCCTCAAACCAACACAGGCAGTGTGTGTATGTGAGCACCTCGTGGGTGAGTATGTGTGGGG  
CACAGGCTGGCTCCCTCAGCTCCCACGTCTTAGAGGGGGCTCCCGAGGAGGTGGAACCTCAACC  
CAGCTCTGCGCAGGAGGGCGGCTGCAGTCTTTTCTCCCTCAAAGGTCTCCGACCCTCAGCTGG  
AGGCGGGCATCTTTCCTAAAGGGTCCCCATAGGGTCTGGTTCCACCCCATCCAGGTCTGTGG  
TCAGAGCCTGGGAGGGTTCCCTACGATGGTTAGGGGTGCCCCATGGAGGGGCTGACTGCCCCA  
CATTGCCTTTCAGACAGGACACGAGCATGAGGTAAGGCCGCCCTGACCTGGACTTCAGGGGGA  
GGGGGTAAAGGGAGAGAGGAGGGGGGGCTAGGGGGTCTCTAGATCAGTGGGGGCACTGCAGGT  
GGGGCTCTCCCTATACCTGGGACACCTGCTGGATGTCACCTCTGCAACCACACCCATGTGGTG  
GTTTCATGAACAGACCACGCTCCTCTGCCTTCTCCTGGCCTGGGACACACAGAGCCACCCCGG  
CCTTGTGAGTGACCCAGAGAAGGGAGGCCTCGGGAGAAGGGGTGCTCGTAAGCCAACACCAGC  
GTGCCGCGGCCTGCACACCCTTCGGACATCCCAGGCACGAGGGTGTCGTGGATGTGGCCACAC  
ATAGGACCACACGTCCCAGCTGGGAGGAGAGGCCTGGGGCCCCCAGGGAGGGAGGCAGGGGGT  
GGGGGACATGGAGAGCTGAGGCAGCCTCGTCTCCCCGCAGCCTGGTATCGCCAGCCTTAAGGT  
GTCTGGAGCCCCCACACTTGGCCAACCTGACCTTGGGAAGATGCTGCTGAGTGTCTCAAGCAGC  
ACTGACAGCAGCTGGGCCTGCCCCAGGGCAACGTGGGGGCGGAGACTCAGCTGGACAGCCCCT  
GCCTGTCACTCTGGAGCTGGGCTGCTGCTGCCTCAGGACCCCCCTCTCCGACCCCGGACAGAGC  
TGAGCTGGCCAGGGCCAGGAGGGCGGGAGGGAGGGAATGGGGGTGGGCTGTGCGCAGCATCAG  
CGCCTGGGCAGGTCCGCAGAGCTGCGGGATGTGATTAAAGTCCCTGATGTTTCTC



374/615

**FIGURE 370**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76399

><subunit 1 of 1, 157 aa, 1 stop

><MW: 17681, pI: 7.65, NX(S/T): 1

MALLLCVLCLTAALAHGCLHCHSNFSKKFSFYRHHVNFKSWWVGDI PVSGALLTDWSDDTMKE

LHLAIPAKITREKLDQVATAVYQMMDQLYQGKMYFPGYFPNELRNIFREQVHLIQNAIIERHL

APGSWGGGQLSREGPSLAPEGSMPSPRGDLP

**Signal peptide:**

amino acids 1-15

375/615

**FIGURE 371**

GCCGGCTGTGCAGAGACGCC**ATG**TACCGGCTCCTGTCAGCAGTGACTGCCCGGGCTGCCGCCC  
CCGGGGGCTTGGCCTCAAGCTGCGGACGACGCGGGGTCCATCAGCGCGCCGGGCTGCCGCCTC  
TCGGCCACGGCTGGGTGCGGGGCTCGGGCTGGGGCTGGCGCTCGGGGTGAAGCTGG  
CAGGTGGGCTGAGGGGCGCGGCCCCGGCGCAGTCCCCCGCGGCCCCGACCTGAGGCGTCGC  
CTCTGGCCGAGCCGCCACAGGAGCAGTCCCTCGCCCCGTGGTCTCCGCAGACCCCGGCGCCGC  
CCTGCTCCAGGTGCTTCGCCAGAGCCATCGAGAGCAGCCGCGACCTGCTGCACAGGATCAAGG  
ATGAGGTGGGCGCACCGGGCATAGTGGTTGGAGTTTCTGTAGATGGAAAAGAAGTCTGGTCAG  
AAGGTTTAGGTTATGCTGATGTTGAGAACCGTGTACCATGTAAACCAGAGACAGTTATGCGAA  
TTGCTAGCATCAGCAAAAAGTCTCACCATGGTTGCTCTTGCCAAATTGTGGGAAGCAGGGAAAC  
TGGATCTTGATATTCCAGTACAACATTATGTTCCCGAATTCCCAGAAAAAGAATATGAAGGTG  
AAAAGGTTTCTGTCACAACAAGATTACTGATTTCCCATTTAAGTGGAATTCGTCATTATGAAA  
AGGACATAAAAAAGGTGAAAGAAGAGAAAGCTTATAAAGCCTTGAAGATGATGAAAGAGAATG  
TTGCATTTGAGCAAGAAAAAGAAGGCAAAAGTAATGAAAAGAATGATTTTACTAAATTTAAAA  
CAGAGCAGGAGAATGAAGCCAAATGCCGGAATTCAAAACCTGGCAAGAAAAAGAATGATTTTG  
AACAAGGCGAATTATATTTGAGAGAAAAGTTGAAAATTC AATTGAATCCCTAAGATTATTTA  
AAAATGATCCTTTGTTCTTCAAACCTGGTAGTCAGTTTTTGTATTCAACTTTTGGCTATACCC  
TACTGGCAGCCATAGTAGAGAGAGCTTCAGGATGTAAATATTTGGACTATATGCAGAAAATAT  
TCCATGACTTGATATGCTGACGACTGTGCAGGAAGAAAACGAGCCAGTGATTTACAATAGAG  
CAAGG**TAA**ATGAATACCTTCTGCTGTGTCTAGCTATATCGCATCTTAACACTATTTTATTAAT  
TAAAAGTCAAATTTTCTTTGTTTCCATTCCAAAATCAACCTGCCACATTTTGGGAGCTTTTCT  
ACATGTCTGTTTTCTCATCTGTAAAGTGAAGGAAGTAAACATGTTTATAAAGTAAAAAAA

376/615

**FIGURE 372**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76522

&gt;&lt;subunit 1 of 1, 373 aa, 1 stop

&gt;&lt;MW: 41221, pI: 8.54, NX(S/T): 0

MYRLLSAVTARAAAPGGLASSCGRRGVHQRAGLPPLGHGWVGGGLGLGLGLALGVKLAGGLRGA  
APAQSPAAPDPEASPLAEPPEQSLAPWSPQTPAPPCSRCFARAIESSRDLLHRIKDEVGAPG  
IVVGVSVDGKEVWSEGLGYADVENRVPCKPETVMRIASISKSLTMVALAKLWEAGKLDLDIPV  
QHYVPEFPEKEYEGEKVSVTTTRLLISHLSGIRHYEKDIKKVKEEKAYKALKMMKENVAFEQEK  
EGKSNEKNDFTKFKTEQENEAKCRNSKPGKKKNDFEQGELYLREKFENSIESLRLFKNDPLFF  
KPGSQFLYSTFGYTLLAAIVERASGCKYLDYMQKIFHDLMLTTVQEENEPVIYNRAR

**Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 39-60

377/615

**FIGURE 373**

GACTACGGGGAGAGAGAGAGAGACCAGGACAGCTGCTGAGACCTCTAAGAAGTCCAGATACTAA  
GAGCAAAG**ATG**TTTCAAACCTGGGGGCCCTCATTGTCTTCTACGGGCTGTTAGCCCAGACCATGG  
CCAGTTTGGAGGCCTGCCCCGTGCCCTGGACCAGACCCTGCCCTTGAATGTGAATCCAGCCCTG  
CCCTTGAGTCCACAGGTCTTGCAGGAAGCTTGACAAATGCCCTCAGCAATGGCCTGCTGTCT  
GGGGGCCTGTTGGGCATTCTGGAAAACCTTCCGCTCCTGGACATCCTGAAGCCTGGAGGAGGT  
ACTTCTGGTGGCCTCCTTGGGGGACTGCTTGGAAGTACGTCAGTGATTCCTGGCCTGAAC  
AACATCATTGACATAAAGGTCAGTACCCCCAGCTGCTGGAACCTTGGCCTTGTGCAGAGCCCT  
GATGGCCACCGTCTCTATGTCACCATCCCTCTCGGCATAAAGCTCCAAGTGAATACGCCCCCTG  
GTCGGTGCAAGTCTGTTGAGGCTGGCTGTGAAGCTGGACATCACTGCAGAAATCTTAGCTGTG  
AGAGATAAGCAGGAGAGGATCCACCTGGTCCTTGGTGACTGCACCCATTCCCCTGGAAGCCTG  
CAAATTTCTCTGCTTGATGGACTTGGCCCCCTCCCCATTCAAGGTCTTCTGGACAGCCTCACA  
GGGATCTTGAATAAAGTCCTGCCTGAGTTGGTTCAGGGCAACGTGTGCCCTCTGGTCAATGAG  
GTTCTCAGAGGCTTGGACATCACCCCTGGTGCATGACATTGTTAACATGCTGATCCACGGACTA  
CAGTTTGTCAATCAAGGTCT**TAA**GCCTTCCAGGAAGGGGCTGGCCTCTGCTGAGCTGCTTCCCAG  
TGCTCACAGATGGCTGGCCCATGTGCTGGAAGATGACACAGTTGCCTTCTCTCCGAGGAACCT  
GCCCCCTCTCCTTTCCCACCAGGCGTGTGTAACATCCCATGTGCCTCACCTAATAAAATGGCT  
CTTCTTATGCA

378/615

**FIGURE 374**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76533  
><subunit 1 of 1, 256 aa, 1 stop  
><MW: 26713, pI: 5.62, NX(S/T): 0  
MFQTGGLIVFYGLLAQTMAQFGGLPVPLDQTLPLNVNPALPLSPTGLAGSLTNALSNGLL  
SGLLGILENLPLLDILKPGGGTSGGLLGGLLGKVTSVIPGLNNIIDIKVTDQPQLELGL  
VQSPDGHRLYVTIPLGIKLQVNTPLVGASLLRLAVKLDITAEILAVRDKQERIHVLGDC  
THSPGSLQISLLDGLGPLPIQGLLDLSTGILNKVLPVLVQGNVCPLVNEVLRGLDITLVH  
DIVNMLIHGLQFVIKV

**Important features of the protein:****Signal peptide:**

Amino acids 1-19

**Transmembrane domain:**

Amino acids 79-97

**N-myristoylation sites:**

Amino acids 46-52;49-55;58-64;62-68;66-72;80-86;81-87;  
82-88;85-91;86-92;89-95;202-208;233-239

379/615

**FIGURE 375**

AGTTCTGAGAAAGAAGGAAATAAACACAGGCACCAAACCACTATCCTAAGTTGACTGTCCTTT  
AAAT**ATGT**CAAGATCCAGACTTTTCAGTGTCACCTCAGCGATCTCAACGATAGGGATCTTGTG  
TTTGCCGCTATTCCAGTTGGTGCTCTCGGACCTACCATGCGAAGAAGATGAAATGTGTGTAA  
TTATAATGACCAACACCCTAATGGCTGGTATATCTGGATCCTCCTGCTGCTGGTTTTGGTGGC  
AGCTCTTCTCTGTGGAGCTGTGGTCCTCTGCCTCCAGTGCTGGCTGAGGAGACCCCGAATTGA  
TTCTCACAGGCGCACCATGGCAGTTTTTGCTGTTGGAGACTTGGACTCTATTTATGGGACAGA  
AGCAGCTGTGAGTCCAACCTGTTGGAATTCACCTTCAAACCTCAAACCCCTGACCTATATCCTGT  
TCCTGCTCCATGTTTTGGCCCTTTAGGCTCCCCACCTCCATATGAAGAAATTGTAAAAACAAC  
**CTGA**TTTTAGGTGTGGATTATCAATTTAAAGTATTAACGACATCTGTAATTCAAAAACATCAA  
ATTTAGGAATAGTTATTTTCAGTTGTTGGAAATGTCCAGAGATCTATTCATATAGTCTGAGGAA  
GGACAATTCGACAAAAGAATGGATGTTGGAAAAAATTTTGGTCATGGAGATGTTTAAATAGTA  
AAGTAGCAGGCTTTTGATGTGTCACTGCTGTATCATACTTTTATGCTACACAACCAAATTAAT  
GCTTCTCCACTAGTATCCAAACAGGCAACAATTAGGTGCTGGAAGTAGTTTCCATCACATTTA  
GGACTCCACTGCAGTATACAGCACACCATTTTCTGCTTTAAACTCTTTCCTAGCATGGGGTCC  
ATAAAAATTATTATAATTTAACAATAGCCCAAGCCGAGAATCCAACATGTCCAGAACCAGAAC  
CAGAAAGATAGTATTTGAATGAAGGTGAGGGGAGAGAGTAGGAAAAAGAAAAGTTTGGAGTTG  
AAGGGTAAAGGATAAATGAAGAGGAAAAGGAAAAGATTACAAGTCTCAGCAAAAACAAGAGGT  
TTTATGCCCCAACCTGAAGAGGAAGAAATTGTAGATAGAAGGTGAAGGAGATTGCTGAAGATA  
TAGAGCACATATAATGCCAACACGGGGAGAAAAGAAAATTTCCCCTTTTACAGTAATGAATGT  
GGCCTCCATAGTCCATAGTGTCTCTGGAGCCTCAGGGCTTGGCATTTATTCAGCATCATG  
CTAAGAACCTTCGGCATAGGTATCTGTTCCCATGAGGACTGCAGAAGTAGCAATGAGACATCT  
TCAAGTGGCATTTTGGCAGTGGCCATCAGCAGGGGGACAGACAAAACATCCATCACAGATGA  
CATATGATCTTCAGCTGACAAATTTGTTGAACAAAACAATAAACATCAATAGATATCTAAAAA

380/615

**FIGURE 376**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77303

&gt;&lt;subunit 1 of 1, 146 aa, 1 stop

&gt;&lt;MW: 16116, pI: 4.99, NX(S/T): 0

MSRSRLFSVTSAISTIGILCLPLFQLVLSDLPCCEEDEMCVNYNDQHPNGWYIWILLLLVLVAA

LLCGAVVLCLQCWLRRPRIDSHRRTMAVFAVGDLDSIYGTEAAVSPTVGIHLQTQTPDLYPVP

APCFGPLGSPPPYEEIVKTT

**Signal peptide:**

amino acids 1-29

**Transmembrane domain:**

amino acids 52-70

381/615

**FIGURE 377**

CGCGGATCGGACCCAAGCAGGTCGGCGGGCGGGCGGCAGGAGAGCGGCCGGGCGTCAGCTCCTCG  
ACCCCCGTGTCGGGCTAGTCCAGCGAGGCGGACGGGCGGGCGTGGGCCC**ATG**GCCAGGCCCGGC  
ATGGAGCGGTGGCGCGACCGGCTGGCGCTGGTGACGGGGGGCCTCGGGGGGCATCGGCGCGGCC  
GTGGCCCGGGCCCTGGTCCAGCAGGGACTGAAGGTGGTGGGCTGCGCCCGCACTGTGGGCAAC  
ATCGAGGAGCTGGCTGCTGAATGTAAGAGTGCAGGCTACCCCGGGACTTTGATCCCCACAGA  
TGTGACCTATCAAATGAAGAGGACATCCTCTCCATGTTCTCAGCTATCCGTTCTCAGCACAGC  
GGTGTAGACATCTGCATCAACAATGCTGGCTTGGCCCGGCCTGACACCCTGCTCTCAGGCAGC  
ACCAGTGGTTGGAAGGACATGTTCAATGTGAACGTGCTGGCCCTCAGCATCTGCACACGGGAA  
GCCTACCAGTCCATGAAGGAGCGGAATGTGGACGATGGGCACATCATTAAACATCAATAGCATG  
TCTGGCCACCGAGTGTTACCCCTGTCTGTGACCCACTTCTATAGTGCCACCAAGTATGCCGTC  
ACTGCGCTGACAGAGGGACTGAGGCAAGAGCTTCGGGAGGCCAGACCCACATCCGAGCCACG  
TGCATCTCTCCAGGTGTGGTGGAGACACAATTCGCCTTCAAACCTCCACGACAAGGACCCTGAG  
AAGGCAGCTGCCACCTATGAGCAAATGAAGTGTCTCAAACCCGAGGATGTGGCCGAGGCTGTT  
ATCTACGTCCTCAGCACCCCCGCACACATCCAGATTGGAGACATCCAGATGAGGCCACCGGAG  
CAGGTGACC**TAG**TGACTGTGGGAGCTCCTCCTTCCCTCCCCACCCTTCATGGCTTGCCTCCTG  
CCTCTGGATTTTAGGTGTTGATTTCTGGATCACGGGATACCACTTCCTGTCCACACCCCGACC  
AGGGGCTAGAAAATTTGTTTGAGATTTTTATATCATCTTGTCAAATTGCTTCAGTTGTAAATG  
TGAAAAATGGGCTGGGAAAGGAGGTGGTGTCCCTAATTGTTTTACTTGTTAACTTGTTCTTG  
TGCCCTGGGCACTTGGCCTTTGTCTGCTCTCAGTGTCTTCCCTTTGACATGGGAAAGGAGTT  
GTGGCCAAAATCCCCATCTTCTTGACCTCAACGTCTGTGGCTCAGGGCTGGGGTGGCAGAGG  
GAGGCCTTCACCTTATATCTGTGTTGTTATCCAGGGCTCCAGACTTCCTCCTCTGCCTGCCCC  
ACTGCACCCTCTCCCCCTTATCTATCTCCTTCTCGGCTCCCCAGCCAGTCTTGGCTTCTTGT  
CCCCTCCTGGGGTCATCCCTCCACTCTGACTCTGACTATGGCAGCAGAACACCAGGGCCTGGC  
CCAGTGGATTTTCATGGTGATCATTAAGAAAGAAAATCGCAACCAAAAAAAAAAAAA



382/615

**FIGURE 378**

MARPGMERWRDRLALVTGASGGIGA A VARALVQOGLKVVG CARTVGNIEELAAECKSAGYPGT  
LIPYRC DLSNEEDILSMFSAIRSQHSGVDICINNAGLARPD TLLSGSTSGWKDMFN VNV LALS  
ICTREAYQSMKERNVDDGHIININSMGHRVLP LSVTHFY SATKYAVTALTEGLRQELREAQT  
HIRATCISPGVVETQFAFKLHDKDPEKAAATYEQMKCLKPEDVAEAVIYVLSTPAHIQIGDIQ  
MRPTEQVT

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**N-myristoylation sites.**amino acids 18-24, 21-27, 22-28, 24-30, 40-46, 90-96, 109-115,  
199-205**Short-chain alcohol dehydrogenase.**

amino acids 30-42, 104-114

383/615

**FIGURE 379**

GAGCGGAGTAAATCTCCACAAGCTGGGAACAAACCTCGTCCCAACTCCCACCCACCGGCGTT  
TCTCCAGCTCGATCTGGAGGCTGCTTCGCCAGTGTGGGACGCAGCTGACGCCCCGCTTATTAGC  
TCTCGCTGCGTCGCCCCGGCTCAGAAGCTCCGTGGCGGGCGGACCGTGACGAGAAGCCCACG  
GCCAGCTCAGTTCTCTTCTACTTTGGGAGAGAGAGAAAGTCAGATGCCCCTTTTAAACTCCCT  
CTTCAAACTCATCTCCTGGGTGACTGAGTTAATAGAGTGGATACAACCTTGCTGAAGATGAA  
GAATATACAATATTGAGGATATTTTTTTCTTTTTTTTTCAAGTCTTGATTGTGGCTTACCT  
CAAGTTACCATTTTTTCAGTCAAGTCTGTTTGTGTTGCTTCTTCAGAAATGTTTTTACAATCTC  
AAGAAAAATATGTCCAGAAATTGAGTTTACTGTTGCTTGATTTGGACTCATTGTTGGGATT  
GATGTTACTGCACTATACTTTCAACAACCAAGACATCAAAGCAGTGTCAAGTTACGTGAGCA  
AATACTAGACTTAAGCAAAAGATATGTTAAAGCTCTAGCAGAGGAAAATAAGAACACAGTGA  
TGTCGAGAACGGTGCTTCTATGGCAGGATATGCGGATCTGAAAAGAACAATTGCTGTCCTTCT  
GGATGACATTTTGCAACGATTGGTGAAGCTGGAGAACAAAGTTGACTATATTGTTGTGAATGG  
CTCAGCAGCCAACACCACCAATGGTACTAGTGGGAATTTGGTGCCAGTAACCACAAAATAAAG  
AACGAATGTCTCGGGCAGTATCAGATAGCAGTTGAAAATCACCTTGCTGCTGCCATCCACTG  
TGGATTATATCCTATGGCAGAAAAGCTTTATAATTGCTGGCTTAGGACAGAGCAATACTTTAC  
AATAAAAGCTCTACACATTTTCAAGGAGTATGCTGGATTGATGGAATCTAATTCTGTACATA  
AAAATTTTAAAGTTATTTGTTTGCTTTCAAGGCAAGTCTGTTCAATGCTGTACTATGTCCTTAA  
AGAGAATTTGGTAACTTGGTTGATGTGGTAAGCAGATAGGTGAGTTTTGTATAAATCTTTTGT  
GTTTGAGATCAAGCTGAAATGAAAACACTGAAAAACATGGATTCAATTTCTATAACACATTTAT  
TTAAGTATATAACACGTTTTTTGGACAAGTGAAGAATGTTAATCATTCTGTCAATTTGTTCTC  
AATAGATGTAAGTGTAGACTACGGCTATTTGAAAAAATGTGCTTATTGTACTATATTTTGT  
ATTCCAATTATGAGCAGAGAAAGGAAATATAATGTTGAAAATAATGTTTTGAAATCATGACCC  
AAAGAATGTATTGATTTGCACTATCCTTCAGAATAACTGAAGGTTAATTATTGTATATTTTAA  
AAAATTACACTTATAAGAGTATAATCTTGAAATGGGTAGCAGCCACTGTCCATTACCTATCGT  
AAACATTGGGGCAATTTAATAACAGCATTAAAATAGTTGTAACTCTAATCTTATACTTATTG  
AAGAATAAAAGATATTTTTATGATGAGAGTAACAATAAAGTATTCATGATTTTTACATACAT  
GAATGTTCAATTTAAAAGTTTAATCCTTTGAGTGTCTATGCTATCAGGAAAGCACATTATTTCC  
ATATTTGGGTAAATTTTGCTTTTATTATATTGGTCTAGGAGGAAGGGACTTTGGAGAATGGAA  
CTCTTGAGGACTTTAGCCAGGTGTATATAATAAAGGTACTTTTGTGCTGCATTAAATGCTTG  
GAAAGTGTTAACATTATATTATATAAGAGTATCCTTTATGAAATTTTGAATTTGTATAACAGA  
TGCATTAGATATTCATTTTATATAATGGCCACTTAAAATAAGAACATTTAAAATATAAACTAT  
GAAGATTGACTATCTTTTCAGGAAAAAAGCTGTATATAGCACAGGGAACCCTAATCTTGGGTA  
ATTCTAGTATAAAAACAAATTATACTTTTATTTAAATTTCCCTTGTAGCAAATCTAATTGCCAC  
ATGGTGCCCTATATTTCATAGTATTTATTCTCTATAGTAACTGCTTAAGTGCAGCTAGCTTCT  
AGATTTAGACTATATAGAATTTAGATATTGTATTGTTTCGTCATTATAATATGCTACCACATGT  
AGCAATAATTACAATATTTTATTAATAAATATGTGAAATATTGTTTCATGAAAGACAGATT  
TCCAAATCTCTCTTCTCTCTGTACTGTCTACCTTTATGTGAAGAAATTAATTATATGCCA  
TTGCCAGGT

384/615

**FIGURE 380**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77648

&gt;&lt;subunit 1 of 1, 140 aa, 1 stop

&gt;&lt;MW: 15668, pI: 10.14, NX(S/T): 5

MFFTISRKNMSQKLSLLLLVFGLIWGLMLLHYTFQQPRHQSSVKLREQILDLSKRYVKALAE

NKNTVDVENGASMAGYADLKRTIAVLLDDILQRLVKLENKVDYIVVNGSAANTTNGTSGNLVP

VTTNKRTNVSGSIR

**Important features of the protein:****Signal peptide:**

amino acids 1-26

385/615

**FIGURE 381**

AACTTCTAC**ATG**GGCCTCCTGCTGCTGGTGCTCTTCCTCAGCCTCCTGCCGGTGGCCTACACC  
ATCATGTCCCTCCACCCCTCCTTTGACTGCGGGCCGTTCAAGGTGCAGAGTCTCAGTTGCCCGG  
GAGCACCTCCCCTCCCGAGGCAGTCTGCTCAGAGGGCCTCGGCCCAGAATTCAGTTCTGGTT  
TCATGCCAGCCTGTAAAAGGCCATGGAACTTTGGGTGAATCACCGATGCCATTTAAGAGGGTT  
TTCTGCCAGGATGGAAATGTTAGGTCGTTCTGTGTCTGCGCTGTTCAATTCAGTAGCCACCAG  
CCACCTGTGGCCGTTGAGTGCTTGAAAT**TGA**GGAACTGAGAAAATTAATTTCTCATGTATTTTT  
CTCATTTATTTATTAATTTTTTAAGTATAGTTGTACATATTTGGGGGTACATGTGATATTTGG  
ATACATGTATACAATATATAATGATCAAATCAGGGTAACTGGGATATCCATCACATCAAACAT  
TTATTTTTTTATTCTTTTTAGACAGAGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTGGTGCC  
ATCTCAGCTTACTGCAACCTCTGCCTGCCAGGTTCAAGCGATTCTCATGCCTCCACCTCCCAA  
GTAGCTGGGACTACAGGCATGCACCACAATGCCCAACTAATTTTTGTATTTTAGTAGAGACG  
GGGTTTTGCCATGTTGCCCAGGCTGGCCTTGAACCTCGGCCTCAAACAATCCACTTGCCTCG  
GCCTCCCAAAGTGTTATGATTACAGGCGTGAGCCACCGTGCCTGGCCTAAACATTTATCTTTT  
CTTTGTGTTGGGAACTTTGAAATTATACAATGAATTATTGTTAACTGTCATCTCCCTGCTGTG  
CTATGGAACACTGGGACTTCTTCCCTCTATCTAACTGTATATTTGTACCAGTTAACCAACCGT  
ACTTCATCCCCACTCCTCTCTATCCTTCCCAACCTCTGATCACCTCATTCTACTCTCTACCTC  
CATGAGATCCACTTTTTTAGCTCCCACATGTGAGTAAGAAAATGCAATATTTGTCTTTCTGTG  
CCTGGCTTATTTCACTTAACATAATGACTTCCTGTTCCATCCATGTTGCTGCAAATGACAGGA  
TTTCGTTCTTAATTTCAATTAAAATAACCACACATGGCAAAA

386/615

**FIGURE 382**

MGLLLLVLFLSLLPVAYTIMSLPPSFDCGPFRCRVSVAREHLPSRGSLLRGPRPRIPVLVSCQ  
PVKGHGTLGESPMPEFKRVFCQDGNVRSFCVCAVHFSSHQPPVAVECLK

**Important features of the protein:**

**Signal peptide:**

amino acids 1-18

**N-myristoylation site.**

amino acids 86-92

**Zinc carboxypeptidases, zinc-binding region 2 signature.**

amino acids 68-79

387/615

**FIGURE 383**

TTCTGAAGTAACGGAAGCTACCTTGTATAAAGACCTCAACACTGCTGACC**ATG**ATCAGCGCAG  
CCTGGAGCATCTTCCTCATCGGGACTAAAATTGGGCTGTTCCCTTCAAGTAGCACCTCTATCAG  
TTATGGCTAAATCCTGTCCATCTGTGTGTCGCTGCGATGCGGGTTTCATTTACTGTAATGATC  
GCTTTCTGACATCCATTCCAACAGGAATACCAGAGGATGCTACAACCTCTCTACCTTCAGAAC  
ACCAAATAAATAATGCTGGGATTCCTTCAGATTTGAAAACTTGCTGAAAGTAGAAAGAATAT  
ACCTATACCACAACAGTTTAGATGAATTTCCCTACCAACCTCCCAAAGTATGTAAAAGAGTTAC  
ATTTGCAAGAAAATAACATAAGGACTATCACTTATGATTCACTTTCAAAAATTCCTTATCTGG  
AAGAATTACATTTAGATGACAACCTCTGTCTCTGCAGTTAGCATAGAAGAGGGAGCATTCGGAG  
ACAGCAACTATCTCCGACTGCTTTTCCCTGTCCCGTAATCACCTTAGCACAAATTCCTGGGGTT  
TGCCAGGACTATAGAAGAACTACGCTTGGATGATAATCGCATATCCACTATTTTCATCACCAT  
CTCTTCAAGGTCTCACTAGTCTAAAACGCCTGGTTCTAGATGGAAACCTGTTGAACAATCATG  
GTTTAGGTGACAAAGTTTTCTTCAACCTAGTTAATTTGACAGAGCTGTCCCTGGTGCGGAATT  
CCCTGACTGCTGCACCAGTAAACCTTCAGGCACAAACCTGAGGAAGCTTTATCTTCAAGATA  
ACCACATCAATCGGGTGCCCCAAATGCTTTTTCTTATCTAAGGCAGCTCTATCGACTGGATA  
TGTCCAATAATAACCTAAGTAATTTACCTCAGGGTATCTTTGATGATTTGGACAATATAACAC  
AACTGATTCTTCGCAACAATCCCTGGTATTGCGGGTGCAAGATGAAATGGGTACGTGACTGGT  
TACAATCACTACCTGTGAAGGTCAACGTGCGTGGGCTCATGTGCCAAGCCCCAGAAAAGGTTT  
GTGGGATGGCTATTAAGGATCTCAATGCAGAACTGTTTGATTGTAAGGACAGTGGGATTGTAA  
GCACCATTCAGATAACCACTGCAATACCCAACACAGTGTATCCTGCCCCAAGGACAGTGGCCAG  
CTCCAGTGACCAAACAGCCAGATATTAAGAACCCCAAGCTCACTAAGGATCAACAAACCACAG  
GGAGTCCCTCAAGAAAAACAATTACAATTACTGTGAAGTCTGTACCTCTGATACCATTCATA  
TCTCTTGAAACTTGCTCTACCTATGACTGCTTTGAGACTCAGCTGGCTTAAACTGGGCCATA  
GCCCCGCATTTGGATCTATAACAGAAACAATTGTAACAGGGGAACGCAGTGAGTACTTGGTCA  
CAGCCCTGGAGCCTGATTCACCCCTATAAAGTATGCATGGTTCCCATGGAAACCAGCAACCTCT  
ACCTATTTGATGAACTCCTGTTTGTATTGAGACTGAACTGCACCCCTTCGAATGTACAACC  
CTACAACCACCCTCAATCGAGAGCAAGAGAAAGAACCTTACAAAAACCCCAATTTACCTTTGG  
CTGCCATCATTTGGTGGGGCTGTGGCCCTGGTTACCATTGCCCTTCTTGCTTTAGTGTGTTGGT  
ATGTTCATAGGAATGGATCGCTCTTCTCAAGGAACCTGTGCATATAGCAAAGGGAGGAGAAGAA  
AGGATGACTATGCAGAAGCTGGCACTAAGAAGGACAACCTCTATCCTGGAAATCAGGGAACTT  
CTTTTCAGATGTTACCAATAAGCAATGAACCCATCTCGAAGGAGGAGTTTGTAAATACACACCA  
TATTTCCCTCCTAATGGAATGAATCTGTACAAAACAATCACAGTGAAAGCAGTAGTAACCGAA  
GCTACAGAGACAGTGGTATTCCAGACTCAGATCACTCACACT**CATGA**TGCTGAAGGACTCACA  
GCAGACTTGTGTTTTGGGTTTTTAAACCTAAGGGAGGTGATGGT

388/615

**FIGURE 384**

MISAAWSIFLIGTKIGLFLQVAPLSVMAKSCPSVCRC DAGFIYCND RFLTSIPTGIPEDATTL  
YLQNNQINNAGIPSDLKNLLKVERIYLYHNSLDEFPTNLPKYVKELHLQENNIRTITYDSLK  
IPYLEELHLDDNSVSAVSIEEGA FRDSNYLRLLFLSRNHLSTIPWGLPRTIEELRLDDNRIST  
ISSPSLQGLTSLKRLVLDGNLLNNHGLGDKVFFNLVNLTELSLVRNSLTAAPVNLPGTNLRKL  
YLQDNHINRVPPNAFSYLRQLYRLDMSNNNLSNLPQGIFDDLDNITQLILRNNPWYCGCKMKW  
VRDWLQSLPVKVNVRGLMCQAPEKVRGMAIKDLNAELFDCKDSGIVSTIQITTAIPNTVYPAQ  
GQWPAPVTKQPDIKNPCLTKDQQTGSPSRKTITITVKSVTSDTIHISWKLALPMTALRLSWL  
KLGHSPAFGSITETIVTGERSEYLVTALEPDSPYKVCMPMETSNLYLFDETPVCJETETAPL  
RMYNPTTTLNREQEKEPYKNPNLPLAAIIGGAVALVTIALLALVCWYVHRNGSLFSRNCAYSK  
GRRRKDDYAEAGTKKDNSILEIRETSFQMLPISNEPISKEEFVIHTIFPPNGMNLYKNNHSES  
SSNRSYRDSGIPDSDHSHS

**Important features of the protein:****Signal peptide:**

amino acids 1-28

**Transmembrane domain:**

amino acids 531-552

**N-glycosylation sites.**

amino acids 226-229, 282-285, 296-299, 555-558, 626-629, 633-636

**Tyrosine kinase phosphorylation site.**

amino acids 515-522

**N-myristoylation sites.**amino acids 12-17, 172-177, 208-213, 359-364, 534-539, 556-561,  
640-645**Amidation site.**

amino acids 567-570

**Leucine zipper pattern.**

amino acids 159-180

**Phospholipase A2 aspartic acid active site.**

amino acids 34-44

389/615

**FIGURE 385**

CCGTCATCCCCCTGCAGCCACCCTTCCCAGAGTCCTTTGCCCAGGCCACCCAGGCTTCTTGG  
CAGCCCTGCCGGGCCACTTGTCTTCA**ATGT**CTGCCAGGGGGAGGTGGGAAGGAGGTGGGAGGAG  
GGCGTGCAGAGGCAGTCTGGGCTTGGCCAGAGCTCAGGGTGCTGAGCGTGTGACCAGCAGTGA  
GCAGAGGCCGGCCATGGCCAGCCTGGGGCTGCTGCTCCTGCTCTTACTGACAGCACTGCCACC  
GCTGTGGTCCTCCTCACTGCCTGGGCTGGACACTGCTGAAAGTAAAGCCACCATTGCAGACCT  
GATCCTGTCTGCGCTGGAGAGAGCCACCGTCTTCCTAGAACAGAGGCTGCCTGAAATCAACCT  
GGATGGCATGGTGGGGGTCCGAGTGCTGGAAGAGCAGCTAAAAAGTGTCCGGGAGAAGTGGGC  
CCAGGAGCCCCCTGCTGCAGCCGCTGAGCCTGCGCGTGGGGATGCTGGGGGAGAAGCTGGAGGC  
TGCCATCCAGAGATCCCTCCACTACCTCAAGCTGAGTGATCCCAAGTACCTAAGAGAGTTCCA  
GCTGACCCTCCAGCCCGGGTTTTGGAAGCTCCACATGCCTGGATCCACACTGATGCCTCCTT  
GGTGTACCCACGTTCCGGCCCCAGGACTCATTCTCAGAGGAGAGAAGTGACGTGTGCCTGGT  
GCAGCTGCTGGGAACCGGGACGGACAGCAGCGAGCCCTGCGGCCTCTCAGACCTCTGCAGGAG  
CCTCATGACCAAGCCCGGCTGCTCAGGCTACTGCCTGTCCACCAACTGCTCTTCTTCCTCTG  
GGCCAGAATGAGGGGATGCACACAGGGACCACTCCAACAGAGCCAGGACTATATCAACCTCTT  
CTGCGCCAACATGATGGACTTGAACCGCAGAGCTGAGGCCATCGGATACGCCTACCCTACCCG  
GGACATCTTCATGGAAAACATCATGTTCTGTGGAATGGGCGGCTTCTCCGACTTCTACAAGCT  
CCGGTGGCTGGAGGCCATTCTCAGCTGGCAGAAACAGCAGGAAGGATGCTTCGGGGAGCCTGA  
TGCTGAAGATGAAGAATTATCTAAAGCTATTCAATATCAGCAGCATTTTTTCGAGGAGAGTGAA  
GAGGCGAGAAAAACAATTTCCAGATTCTCGCTCTGTTGCTCAGGCTGGAGTACAGTGGCGCAA  
TCTCGGCTCACTGCAACCTTTGCCTCCTGGGTTCAAGCAATTCTCTTGCCTCATCCTCCCGAG  
TAGCTGGGACTACAGGAGCGTGCCACCATACTGGCTAATTTTTATATTTTTTAGTAGAGAC  
AGGGTTTCATCATGTGTGCTCATGCTGGTCTCGAACTCCTGATCTCAAGAGATCCGCCCACCTC  
AGGCTCCCAAAGTGTGGGATTAT**TAG**GTGTGAGCCACCGTGTCTGGCTGAAAAGCACTTTCAA  
GAGACTGTGTTGAATAAAGGGCCAAGGTTCTTGCCACCCAGCACTCATGGGGGCTCTCTCCCC  
TAGATGGCTGCTCCTCCCACAACACAGCCACAGCAGTGGCAGCCCTGGGTGGCTTCCTATACA  
TCCTGGCAGAATAACCCCCAGCAAACAGAGAGCCACACCCATCCACACCGCCACCACCAAGCA  
GCCGCTGAGACGGACGGTTCCATGCCAGCTGCCTGGAGGAGGAACAGACCCCTTTAGTCCTCA  
TCCCTTAGATCCTGGAGGGCACGGATCACATCCTGGGAAGAAGGCATCTGGAGGATAAGCAAA  
GCCACCCCGACACCCAATCTTGGAAGCCCTGAGTAGGCAGGGCCAGGGTAGGTGGGGGCCGGG  
AGGGACCCAGGTGTGAACGGATGAATAAAGTTCAACTGCAACTGAAAAAAAAA



390/615

**FIGURE 386**

MSARGRWEGGRRACRGSGLARAQGAERVTSSEQRPAMASLGLLLLLLLLLTALPPLWSSSLPG  
LDTAESKATIADLILSALERATVFLEQRLPEINLDGMVGVRVLEEQLKSVREKWAQEPLLQPL  
SLRVGMLGEKLEAAIQRS LHYLKLSDPKYLREFQLTLQPGFWKLPHAWIHTDASLVYPTFGPQ  
DSFSEERSDVCLVQLLGTGTDSSSEPCGLSDLCRSLMTKPGCSGYCLSHQLLFFLWARMRGCTQ  
GPLQQSQDYINLFCANMMDLNRRAEAIGYAYPTRDIFMENIMFCGMGGFSDFYKLRWLEAILS  
WQKQQEGCGEPDAEDEELSKAIQYQQHFSRRVKRREKQFPDSRSVAQAGVQWRNLGSLQPLP  
PGFKQFSCILLPSSWDYRSVPPYLANFYIFLVETGFHHVAHAGLELLISRDPPTSGSQSVGL

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 39-56

**Tyrosine kinase phosphorylation sites.**

amino acids 149-156, 274-282

**N-myristoylation sites.**

amino acids 10-16, 20-26, 63-69, 208-214

**Amidation site.**

amino acids 10-14

**Glycoprotein hormones beta chain signature 1.**

amino acids 230-237

391/615

**FIGURE 387**

GGTCTGAGTGCAGAGCTGCTGTCAATGGCGGCCGCTCTGTGGGGCTTCTTTCCCGTCCTGCTGC  
TGCTGCTGCTATCGGGGGATGTCCAGAGCTCGGAGGTGCCCGGGGCTGCTGCTGAGGGATCGG  
GAGGGAGTGGGGTCCGGCATAGGAGATCGCTTCAAGATTGAGGGGCGTGCAAGTTGTTCCAGGGG  
TGAAGCCTCAGGACTGGATCTCGGCGGCCCGAGTGCTGGTAGACGGAGAAGAGCACGTCGGTT  
TCCTTAAGACAGATGGGAGTTTTGTGGTTCATGATATACCTTCTGGATCTTATGTAGTGGAAG  
TTGTATCTCCAGCTTACAGATTTGATCCCGTTCGAGTGGATATCACTTCGAAAGGAAAAATGA  
GAGCAAGATATGTGAATTACATCAAAACATCAGAGGTTGTCAGACTGCCCTATCCTCTCCAAA  
TGAAATCTTCAGGTCCACCTTCTTACTTTATTAAGGGAATCGTGGGGCTGGACAGACTTTC  
TAATGAACCCAATGGTTATGATGATGGTTCTTCCTTTATTGATATTTGTGCTTCTGCCTAAAG  
TGGTCAACACAAGTGATCCTGACATGAGACGGGAAATGGAGCAGTCAATGAATATGCTGAATT  
CCAACCATGAGTTGCCTGATGTTTCTGAGTTCATGACAAGACTCTTCTCTTCAAAATCATCTG  
GCAAATCTAGCAGCGGCAGCAGTAAACAGGCAAAAGTGGGGCTGGCAAAGGAGGTAGTCAG  
GCCGTCCAGAGCTGGCATTTCACAAAACACGGCAACACTGGGTGGCATCCAAGTCTTGAAAA  
CCGTGTGAAGCAACTACTATAAACTTGAGTCATCCCGACGTTGATCTCTTACAAGTGTGTATGTT  
AACTTTTTAGCACATGTTTTGTACTTGGTACACGAGAAAACCCAGCTTTCATCTTTTGTCTGT  
ATGAGGTCAATATTGATGTCACTGAATTAATTACAGTGTCTATAGAAAATGCCATTAATAAA  
TTATATGAAGTACTATACATTATGTATATTAATTAACATCTTAATCCAGAAATCAAAAAA  
AAAAAAAAAAAAAAAAAAAAA

392/615

**FIGURE 388**

MAAALWGFFPVLLLLLLLLSGDVQSSEVPGAAAEGSGGSGVGIGDRFKIEGRAVVPGVKPDWIS  
AARVLVDGEEHVGFLKTDGSFVVDIPSGSYVVEVVSPAYRFDVPRVDITSKGKMRARYVNYI  
KTSEVVRLPYPQLQMKSSGPPSYFIKRESWGWTDFLMNPMVMMMLVPLLI FVLLPKVVNTSDPD  
MRREMEQSMNMLNSNHELPDVSEFMTRLFSSKSSGKSSSGSSKTGKSGAGKRR

**Important features of the protein:****Signal sequence:**

amino acids 1-23

**Transmembrane domain:**

amino acids 161-182

**N-glycosylation site.**

amino acids 184-187

**Glycosaminoglycan attachment sites.**

amino acids 37-40, 236-239

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 151-154

**N-myristoylation sites.**

amino acids 33-38, 36-41, 38-44, 229-234

**Amidation site.**

amino acids 238-241

**ATP/GTP-binding site motif A (P-loop).**

amino acids 229-236

393/615

**FIGURE 389**

GTCTGTGCTTGGAGGAAGCCGCGGAACCCCCAGCGTCCGTCC**ATG**GCGTGGAGCCTTGGGAG  
CTGGCTGGGTGGCTGCCTGCTGGTGTGAGCATTTGGGAATGGTACCACCTCCCGAAAATGTCAG  
AATGAATTCTGTTAATTTCAAGAACATTCTACAGTGGGAGTCACCTGCTTTTGCCAAAGGGAA  
CCTGACTTTTACAGCTCAGTACCTAAGTTATAGGATATTCCAAGATAAATGCATGAATACTAC  
CTTGACGGAATGTGATTTTCTCAAGTCTTTCCAAGTATGGTGACCACACCTTGAGAGTCAGGGC  
TGAATTTGCAGATGAGCATTGAGACTGGGTAAACATCACCTTCTGTCCTGTGGATGACACCAT  
TATTGGACCCCCTGGAATGCAAGTAGAAGTACTTGCTGATTCTTTACATATGCGTTTCTTAGC  
CCCTAAAATTGAGAATGAATACGAACTTGGAATGATAAGAAATGTGTATAACTCATGGACTTA  
TAATGTGCAATACTGGAAAAACGGTACTGATGAAAAGTTTCAAATTACTCCCCAGTATGACTT  
TGAGGTCTCAGAAACCTGGAGCCATGGACAACCTTATTGTGTTCAAGTTCGAGGGTTTCTTCC  
TGATCGGAACAAAGCTGGGGAATGGAGTGAGCCTGTCTGTGAGCAAACAACCCATGACGAAAC  
GGTCCCCTCCTGGATGGTGGCCGTATCCTCATGGCCTCGGTCTTCATGGTCTGCCTGGCACT  
CCTCGGCTGCTTCTCCTTGCTGTGGTGCCTTTACAAGAAGACAAAGTACGCCCTTCTCCCCTAG  
GAATTCTCTTCCACAGCACCTGAAAGAGTTTTTGGGCCATCCTCATCATAACACACTTCTGTT  
TTTCTCCTTTCCATTGTCTGGATGAGAATGATGTTTTTGACAAGCTAAGTGTGATTGCAGAAGA  
CTCTGAGAGCGGCAAGCAGAATCCTGGTGACAGCTGCAGCCTCGGGACCCCGCCTGGGCAGGG  
GCCCCAAAGCT**TAG**GCTCTGAGAAGGAAACACACTCGGCTGGGCACAGTGACGTACTCCATCTC  
ACATCTGCCCTCAGTGAGGGATCAGGGCAGCAAACAAGGGCCAAGACCATCTGAGCCAGCCCCA  
CATCTAGAACTCCAGACCTGGAATAGCCACCAGAGAGCTACATTTTAAAGGCTGTCTTGGCA  
AAAATACTCCATTTGGGAACTCACTGCCTTATAAAGGCTTTCATGATGTTTTCAGAAGTTGGC  
CACTGAGAGTGTAATTTTCAGCCTTTTATATCACTAAAATAAGATCATGTTTTAATTGTGAGA  
AACAGGGCCGAGCACAGTGGCTCACGCCTGTAATACCAGCACCTTAGAGGTCGAGGCAGGCGG  
ATCACTTGAGGTGAGGAGTTCAAGACCAGCCTGGCCAATATGGTGAAACCCAGTCTCTACTAA  
AAATACAAAAATTAGCTAGGCATGATGGCGCATGCCTATAATCCCAGCTACTCGAGTGCCTGA  
GGCAGGAGAATTGCATGAACCCGGGAGGAGGAGGAGGAGGTTGCAGTGAGCCGAGATAGCGGC  
ACTGCACTCCAGCCTGGGTGACAAAGTGAGACTCCATCTCAAAAAAAAAAAAAAAAAAATTGTG  
AGAAACAGAAATACTTAAATGAGGAATAAGAATGGAGATGTTACATCTGGTAGATGTAACAT  
TCTACCAGATTATGGATGGACTGATCTGAAAATCGACCTCAACTCAAGGGTGGTCAGCTCAAT  
GCTACACAGAGCACGGACTTTTGGATTCTTGCAGTACTTTGAATTTATTTTCTACCTATAT  
ATGTTTTATATGCTGCTGGTGTCCATTAAAGTTTTACTCTGTGTG

394/615

**FIGURE 390**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83551  
><subunit 1 of 1, 325 aa, 1 stop  
><MW: 37011, pI: 5.09, NX(S/T): 4  
MAWSLGSWLGGCLLVSALGMVPPPENVRMNSVNFKNILQWESPAFAKGNLTFTAQYLSYRIFQ  
DKCMNTTLTECDFSSLSKYGDHTLRVRAEFADSHSDWVNITFCPVDDTIIGPPGMQVEVLADS  
LHMRFLAPKIENEYETWTMKNVYNSWTYNVQYWKNGTDEKQITPOYDFEVLRLNLEPWTTYCV  
QVRGFLPDRNKAGEWSEPVCETTHDETVPSSWMVAVILMASVFMVCLALLGCFSLLWCYKKT  
KYAFSPRNSLPQHLKEFLGHPHNTLLFFSFPLSDENDVFDKLSVIAEDSESGKQNP GDSCSL  
GTPPGQGPQS

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 222-245

**N-glycosylation sites.**

amino acids 49-53, 68-72, 102-106, 161-165

**N-myristoylation sites.**

amino acids 6-12, 316-322

395/615

**FIGURE 391**

CTGTGCAGCTCGAGGCTCCAGAGGCACACTCCAGAGAGAGCCAAGGTTCTGACGCG**ATG**AGGA  
AGCACCTGAGCTGGTGGTGGCTGGCCACTGTCTGCATGCTGCTCTTCAGCCACCTCTCTGCGG  
TCCAGACGAGGGGCATCAAGCACAGAATCAAGTGGAACCGGAAGGCCCTGCCAGCACTGCCC  
AGATCACTGAGGCCCAGGTGGCTGAGAACCGCCCGGGAGCCTTCATCAAGCAAGGCCGCAAGC  
TCGACATTGACTTCGGAGCCGAGGGCAACAGGTACTACGAGGCCAACTACTGGCAGTTCCCCG  
ATGGCATCCACTACAACGGCTGCTCTGAGGCTAATGTGACCAAGGAGGCATTTGTCACCGGCT  
GCATCAATGCCACCCAGGCGGCGAACCAGGGGGAGTTCCAGAAGCCAGACAACAAGCTCCACC  
AGCAGGTGCTCTGGCGGCTGGTCCAGGAGCTCTGCTCCCTCAAGCATTGCGAGTTTTGGTTGG  
AGAGGGGCGCAGGACTTCGGGTCACCATGCACCAGCCAGTGCTCCTCTGCCTTCTGGCTTTGA  
TCTGGCTCATGGTGAA**ATA**AGCTTGCCAGGAGGCTGGCAGTACAGAGCGCAGCAGCGAGCAAA  
TCCTGGCAAGTGACCCAGCTCTTCTCCCCCAAACCCACGCGTGTTCTGAAGGTGCCCAGGAGC  
GGCGATGCACTCGCACTGCAAATGCCGCTCCCACGTATGCGCCCTGGTATGTGCCTGCGTTCT  
GATAGATGGGGACTGTGGCTTCTCCGTCACTCCATTCTCAGCCCCTAGCAGAGCGTCTGGCA  
CACTAGATTAGTAGTAAATGCTTGATGAGAAGAACACATCAGGCACTGCGCCACCTGCTTCAC  
AGTACTTCCCAACAACCTCTTAGAGGTAGGTGTATTCCCGTTTTACAGATAAGGAACTGAGGC  
CCAGAGAGCTGAAGTACTGCACCCAGCATCACCAGCTAGAAAGTGGCAGAGCCAGGATTCAAC  
CCTGGCTTGTCTAACCCAGGTTTTCTGCTCTGTCCAATTCCAGAGCTGTCTGGTGATCACTT  
TATGTCTCACAGGGACCCACATCCAAACATGTATCTCTAATGAAATTGTGAAAGCTCCATGTT  
TAGAAATAAATGAAAACACCTGA

396/615

## **FIGURE 392**

MRKHLSSWWLATVCMLLFSHLSAVQTRGIKHRIKWNRKALPSTAQITEAQVAENRPGAFIKQG  
RKLDIDFGAEGNRYYEANYWQFPDGIHYNGCSEANVTKEAFVTGCINATQAANQGEFQKPDNK  
LHQQVLWRLVQELCSLKHCEFWLERGAGLRVTMHQPVLLCLLALIWLMLVK

### **Important features of the protein:**

#### **Signal peptide:**

amino acids 1-26

#### **Transmembrane domain:**

amino acids 157-171

#### **N-glycosylation sites.**

amino acids 98-102, 110-114

#### **Tyrosine kinase phosphorylation site.**

amino acids 76-83

#### **N-myristoylation sites.**

amino acids 71-77, 88-94, 93-99, 107-113, 154-160

#### **Amidation site.**

amino acids 62-66

397/615

**FIGURE 393**

TGAAATGACTTCCACGGCTGGGACGGGAACCTTCCACCCACAGCTATGCCTCTGATTGGTGAA  
TGGTGAAGGTGCCTGTCTAACTTTTCTGTAAAAAGAACCAGCTGCCTCCAGGCAGCCAGCCCT  
CAAGCATCACTTACAGGACCAGAGGGACAAGACATGACTGTGATGAGGAGCTGCTTTCGCCAA  
TTTAACACCAAGAAGAATTGAGGCTGCTTGGGAGGAAGGCCAGGAGGAACACGAGACTGAGAG  
**ATGA**ATTTTCAACAGAGGCTGCAAAGCCTGTGGACTTTAGCCAGACCCTTCTGCCCTCCTTTG  
CTGGCGACAGCCTCTCAAATGCAGATGGTTGTGCTCCCTTGCCTGGGTTTTACCCTGCTTCTC  
TGGAGCCAGGTATCAGGGGCCCAGGGCCAAGAATTCCACTTTGGGCCCTGCCAAGTGAAGGGG  
GTTGTTCCCCAGAACTGTGGGAAGCCTTCTGGGCTGTGAAAGACACTATGCAAGCTCAGGAT  
AACATCACGAGTGCCCGGCTGCTGCAGCAGGAGGTTCTGCAGAACGTCTCGGATGCTGAGAGC  
TGTTACCTTGTCCACACCCTGCTGGAGTTCTACTTGA AAACTGTTTTCAAAAACCACCACAAT  
AGAACAGTTGAAGTCAGGACTCTGAAGTCATTCTCTACTCTGGCCAACA ACTTTGTTCTCATC  
GTGTCACAAC TGCAACCCAGTCAAGAAAATGAGATGTTTTCCATCAGAGACAGTGCACACAGG  
CGGTTTTCTGCTATTCCGGAGAGCATTCAAACAGTTGGACGTAGAAGCAGCTCTGACCAAAGCC  
CTTGGGGAAGTGGACATTCTTCTGACCTGGATGCAGAAATCTACAAGCT**CTGA**ATGTCTAGA  
CCAGGACCTCCCTCCCCCTGGCACTGGTTTTGTTCCCTGTGTCATTTCAAACAGTCTCCCTTCC  
TATGCTGTTCACTGGACACTTCACGCCCTTGGCCATGGGTCCCATTCTTGGCCCAGGATTATT  
GTCAAAGAAGTCATTCTTTAAGCAGCGCCAGTGACAGTCAGGGAAGGTGCCTCTGGATGCTGT  
GAAGAGTCTACAGAGAAGATTCTTGTATTTATTACA ACTCTATTTAATTAATGTCAGTATTTT  
AACTGAAGTTCTATTTATTTGTGAGACTGTAAGTTACATGAAGGCAGCAGAATATTGTGCCCC  
ATGCTTCTTTACCCCTCACAATCCTTGCCACAGTGTGGGGCAGTGGATGGGTGCTTAGTAAGT  
ACTTAATAAACTGTGGTGCTTTTTTTTGGCCTGTCTTTGGATTGTTAAAAAACAGAGAGGGATG  
CTTGGATGTAAAACTGAAC TTCAGAGCATGAAAATCACACTGTCTTCTGATATCTGCAGGGAC  
AGAGCATTGGGGTGGGGGTAAGGTGCATCTGTTTGAAAAGTAAACGATAAAATGTGGATTAAA  
GTGCCCAGCACAAAGCAGATCCTCAATAAACATTTTCA TTTCCCACCCACACTCGCCAGCTCAC  
CCCATCATCCCTTTCCCTTGGTGCCCTCCTTTTTTTTTTTATCCTAGTCATTCTTCCCTAATCT  
TCCACTTGAGTGTCAAGCTGACCTTGCTGATGGTGACATTGCACCTGGATGTACTATCCAATC  
TGTGATGACATTCCCTGCTAATAAAAGACAACATAACTCCAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAA



398/615

**FIGURE 394**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA88002

&gt;&lt;subunit 1 of 1, 206 aa, 1 stop

&gt;&lt;MW: 23799, pI: 9.12, NX(S/T): 3

MNFQQRLQSLWTLARPFPCPLLATASQMQMVVLPCLGFTLLLSQVSGAQGQEFHFGPCQVKG  
VVPQKLWEAFWAVKDTMQAQDNITSARLLQQEVLQNVSDAESCYLEVHTLLEFYLKTVFKNHHN  
RTVEVRTLKSFSTLANNFVLIVSQLQPSQENEMFSIRDSAHRRFLLFRRAFKQLDVEAALTKA  
LGEVDILLTWMQKFYKL

**Signal sequence:**

amino acids 1-42

**N-glycosylation sites.**

amino acids 85-89, 99-103, 126-130

399/615

**FIGURE 395**

GCCTTGGCCTCCCAAAGGGCTGGGATTATAGGCGTGACCACCATGTCTGGTCCAGAGTCTCAT  
TTCCTGATGATTTATAGACTCAAAGAAAAC**TATG**TTCAGAAGCTCTCTTCTCTTCTGGCCTC  
CTCTCTGTCTTCTTTCCCTCTTTCTTCTTATTTTAATTAGTAGCATCTACTCAGAGTCATGCA  
AGCTGGAAATCTTTCATTTTGCTTGTGAGTGGGGTAGGTCAGTCTTAGTTTTTATTTTT  
TGAAATTTCAACTTTCAGATTCAGGGGGTACATGTGAAGGTTTGTTTTATGAGTATATTGCAT  
**GAT**GCTGAGGTTTGGGGT

400/615

**FIGURE 396**

MFRSSLLFWPPLCLLSLFLILISSIYSESKLEIFHFACQWGRSLSLSFYFLKFQLSDSGGT  
CEGLFYEYIA

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**N-myristoylation site.**

amino acids 62-68

401/615

**FIGURE 397**

**CATG**CCGCTGCCGCCGCTGCTGCTGTTGCTCCTGGCGGCGCCTTGGGGACGGGCAGTTCCCTG  
TGTCTCTGGTGGTTTTGCCTAAACCTGCAAACATCACCTTCTTATCCATCAACATGAAGAATGT  
CCTACAATGGACTCCACCAGAGGGTCTTCAAGGAGTTAAAGTTACTTACACTGTGCAGTATTT  
CATATATGGGCAAAAGAAATGGCTGAATAAATCAGAATGCAGAAATATCAATAGAACCTACTG  
TGATCTTTCTGCTGAAACTTCTGACTACGAACACCAGTATTATGCCAAAGTTAAGGCCATTTG  
GGGAACAAAGTGTTCCAAATGGGCTGAAAGTGGACGGTTCATCCTTTTTTAGAAACACAAAT  
TGGCCCACCAGAGGTGGCACTGACTACAGATGAGAAGTCCATTTCTGTTGTCCTGACAGCTCC  
AGAGAAGTGGAAGAGAAATCCAGAAGACCTTCCTGTTTCCATGCAACAAATATACTCCAATCT  
GAAGTATAACGTGTCTGTGTTGAATACTAAATCAAACAGAACGTGGTCCCAGTGTGTGACCAA  
CCACACGCTGGTGCTCACCTGGCTGGAGCCGAACACTCTTTACTGCGTACACGTGGAGTCCTT  
CGTCCCAGGGCCCCCTCGCCGTGCTCAGCCTTCTGAGAAGCAGTGTGCCAGGACTTTGAAAGA  
TCAATCATCAGAGTTCAAGGCTAAATCATCTTCTGGTATGTTTTGCCCATATCTATTACCGT  
GTTTCTTTTTTCTGTGATGGGCTATTCCATCTACCGATATATCCACGTTGGCAAAGAGAAACA  
CCCAGCAAATTTGATTTTGATTTATGGAAATGAATTTGACAAAAGATTCTTTGTGCCTGCTGA  
AAAAATCGTGATTAACCTTTATCACCTCAATATCTCGGATGATTCTAAATTTCTCATCAGGA  
TATGAGTTTACTGGGAAAAAGCAGTGATGTATCCAGCCTTAATGATCCTCAGCCCAGCGGGAA  
CCTGAGGCCCCCTCAGGAGGAAGAGGAGGTGAAACATTTAGGGTATGCTTCGCATTTGATGGA  
AATTTTTTGTGACTCTGAAGAAAACACGGAAGGTACTTCTCTCAGCCAGCAAGAGTCCCTCAG  
CAGAACAATACCCCCGGATAAAACAGTCATTGAATATGAATATGATGTCAGAACCACTGACAT  
TTGTGCGGGGCCTGAAGAGCAGGAGCTCAGTTTGCAGGAGGAGGTGTCCACACAAGGAACATT  
ATTGGAGTCGCAGGCAGCGTTGGCAGTCTTGGGCCCCGCAAACGTTACAGTACTCATACACCCC  
TCAGCTCCAAGACTTAGACCCCCCTGGCGCAGGAGCACACAGACTCGGAGGAGGGGCCGGAGGA  
AGAGCCATCGACGACCCTGGTCGACTGGGATCCCCAACTGGCAGGCTGTGTATTTCCTTCGCT  
GTCCAGCTTCGACCAGGATTCAGAGGGCTGCGAGCCTTCTGAGGGGGATGGGCTCGGAGAGGA  
GGGTCTTCTATCTAGACTCTATGAGGAGCCGGCTCCAGACAGGCCACCAGGAGAAAATGAAAC  
CTATCTCATGCAATTCATGGAGGAATGGGGGTTATATGTGCAGATGGAAAAC**TGA**TGCCAACA  
CTTCCTTTTGCCTTTTGTTTCCTGTGCAAACAAGTGAGTCACCCCTTTGATCCCAGCCATAAA  
GTACCTGGGATGAAAGAAGTTTTTTCCAGTTTGTCTAGTGTCTGTGAGAA

402/615

**FIGURE 398**

MPLPPLLLLLLLAAPWGRAVPCVSGGLPKPANITFLSINMKNVLQWTPPEGLQGKVTYTVQYF  
IYGQKKWLNKSECRNINRTYCDLSAETSDYEHQYYAKVKAIWGTKCSKWAESGRFYPPLETQI  
GPPEVALTTDEKSISVVLTAPEKWKRNPEDLPVSMQQIYSNLKYNVSVLNTKSNRTWSQCVTN  
HTLVLTWLEPNTLYCVHVESFVPGPPRAQPSEKQCARTLKDQSSEFKAKIIFWYVLPISITV  
FLFSVMGYSIYRYIHVGKEKHHPANLILYGNFEDKRFFVPAEKIVINFITLNISSDDSKISHQD  
MSLLGKSSDVSSLNDPQPSGNLRPPQEEEEVKHLGYASHLMEIFCDSEENTEGTSLTQQESLS  
RTIPDPKTVIEYEYDVRTTDICAGPEEQELSLQEEVSTQGTLLSQALAVLGPQTLQYSYTP  
QLQDLDPQAQEHDTSEEGPEEPPSTTLVDWDPQTGRLCIPSLSSFDQDSEGCEPSEGDGLGEE  
GLLSRLYEPPAPDRPPGENETYLMQFMEEWGLYVQMEN

**Signal sequence:**

amino acids 1-18

**Transmembrane domain:**

amino acids 240-260

**N-glycosylation sites.**amino acids 31-34, 72-75, 80-83, 171-174, 180-183, 189-192,  
304-307, 523-526**Tyrosine kinase phosphorylation site.**

amino acids 385-392, 518-526

**N-myristoylation sites.**

amino acids 53-58, 106-111, 368-373, 492-497

**Tissue factor**

amino acids 1-278

403/615

**FIGURE 399**

CCGGCG**ATG**TCGCTCGTGCTGCTAAGCCTGGCCGCGCTGTGCAGGAGCGCCGTACCCCGAGAG  
CCGACCGTTCAATGTGGCTCTGAAACTGGGCCATCTCCAGAGTGGATGCTACAACATGATCTA  
ATCCCCGGAGACTTGAGGGACCTCCGAGTAGAACCTGTTACAAC TAGTGTTGCAACAGGGGAC  
TATTC AATTTTGATGAATGTAAGCTGGGTACTCCGGGCAGATGCCAGCATCCGCTTGTTGAAG  
GCCACCAAGATTTGTGTGACGGGCAAAAGCAACTTCCAGTCCTACAGCTGTGTGAGGTGCAAT  
TACACAGAGGCCTTCCAGACTCAGACCAGACCCTCTGGTGGTAAATGGACATTTTCCTACATC  
GGCTTCCCTGTAGAGCTGAACACAGTCTATTTTCATTGGGGCCCATAATATTCCTAATGCAAAT  
ATGAATGAAGATGGCCCTTCCATGTCTGTGAATTTACCTCACCAGGCTGCCTAGACCACATA  
ATGAAATATAAAAAAAGTGTGTCAAGGCCGGAAGCCTGTGGGATCCGAACATCACTGCTTGT  
AAGAAGAATGAGGAGACAGTAGAAGTGAACCTCACAACCACTCCCCTGGGAAACAGATACATG  
GCTCTTATCCAACACAGCACTATCATCGGGTTTTCTCAGGTGTTTGAGCCACACCAGAAGAAA  
CAAACGCGAGCTTCAGTGGTGATTCCAGTGACTGGGGATAGTGAAGGTGCTACGGTGCAGCTG  
ACTCCATATTTTCCTACTTGTGGCAGCGACTGCATCCGACATAAAGGAACAGTTGTGCTCTGC  
CCACAAACAGGCGTCCCTTTCCCTCTGGATAACAACAAAAGCAAGCCGGGAGGCTGGCTGCCT  
CTCCTCCTGCTGTCTCTGCTGGTGGCCACATGGGTGCTGGTGGCAGGGATCTATCTAATGTGG  
AGGCACGAAAGGATCAAGAAGACTTCCTTTTCTACCACCACACTACTGCCCCCATTAAAGTT  
CTTGTGGTTTACCCATCTGAAATATGTTTCCATCACACAATTTGTTACTTCACTGAATTTCTT  
CAAAACCATTCAGAAAGTGAGGTCATCCTTGAAAAGTGGCAGAAAAAGAAAATAGCAGAGATG  
GGTCCAGTGCAGTGGCTTGCCACTCAAAGAAGGCAGCAGACAAAGTCGTCTTCCTTCTTTCC  
AATGACGTCAACAGTGTGTGCGATGGTACCTGTGGCAAGAGCGAGGGCAGTCCCAGTGAGAAC  
TCTCAAGACCTCTTCCCCCTTGCCTTTAACCTTTTCTGCAGTGATCTAAGAAGCCAGATT CAT  
CTGCACAAATACGTGGTGGTCTACTTTAGAGAGATTGATACAAAAGACGATTACAATGCTCTC  
AGTGTCTGCCCCAAGTACCACCTCATGAAGGATGCCACTGCTTTCTGTGCAGAACTTCTCCAT  
GTCAAGCAGCAGGTGTCAGCAGGAAAAAGATCACAAGCCTGCCACGATGGCTGCTGCTCCTTG  
**TAG**

404/615

**FIGURE 400**

MSLVLLSLAALCRSAVPREPTVQCGSETGPSPEWMLQHDLI PGDLRDLRVEPVTT SVATGDYS  
ILMNVS WVL RADASIRLLKATKICVTGKS NFQSYSCVRCNYTEAFQTQTRPSGGKWTFSYIGF  
PVELNTVYFIGAHNIPNANMNEDGPSMSVNFTSPGC LDHIMKYKKKCVKAGSLWDPNITACKK  
NEETVEVNFTTTPLGNRYMALIQHSTIIGFSQVFEPHQKKQTRASVVIPVTGDSEGATVQLTP  
YFPTCGSDCIRHKGTVVLCPQTGVFPFLDNNKSKPGGWLPLLLLSLLVATWVLVAGIYLMWRH  
ERIKKTSFSTTTLLPPIKVLVVYPSEICFHHTICYFTEFLQNHCRSEVILEKWQKKKIAEMGP  
VQWLATQKKAADKVVFLSNDVNSVCDGTCGKSEGPSSENSQDLFPLAFNLFCSDLRSQIHLH  
KYVVVYFREIDTKDDYNALSVCPKYHLMKDATAFCAELLHV KQQVSAGKRSQACHDGCCSL

**Important features of the protein:****Signal peptide:**

amino acids 1-14

**Transmembrane domain:**

amino acids 290-309

**N-glycosylation sites.**amino acids 67 - 71, 103 - 107, 156 - 160, 183 - 187, 197 - 201  
and 283 - 287**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 228 - 232 and 319 - 323

**Casein kinase II phosphorylation sites.**

amino acids 178 - 182, 402 - 406, 414 - 418 and 453 - 457

**N-myristoylation site.**

amino acids 116-122

**Amidation site.**

amino acids 488-452

405/615

**FIGURE 401**

GGGAACAGGGAACTATCAGCCCGTCGGCCTCCGGGCCCTGCATTCTCTAGCC**ATGG**ACCG  
GGACCTTTTGCGGCAGTCGCTAAATTGCCACGGGTCGTCTTTGCTCTCTCTACTTTCGGAG  
CGAACAGCAGGACAATCCACACTTCCGTAGCCTCCTGGGGTCGGCCGCCGAGCCAGCCCCG  
GGGCCCCGCCGCCAGCACCCGTTGCAGGGCAGAAAAGAGAAGAGAGTTGACAACATCGA  
GATACAGAAATTATCTCTCCAAAAAAGCGGATCTGCTTTTTGCACTTTCCTGAAATCAGA  
TGCACCTGCAACTTCTGAAATTAATGAAGACAGTGAAGATCATTATGCAATCATGCCACC  
TTTAGAGCAATTATGAGATACCTAGTATGGATCGGAGAGAGCTGTTTTTCCGAGATAT  
TGAGCGTGGTGATATAGTGATTGGAAGAATTAGTTCTATTTCGGAATTCGGTTTTTTCAT  
GGTGTGATCTGTTTAGGAAGTGGTATCATGAGAGATATAGCCCACTTAGAAATCACAGC  
TCTTTGTCCCTTAAGAGATGTGCCCTTCTCACAGTAACCATGGGGATCCTTTATCATATTA  
CCAAACTGGTGACATCATTCGAGCTGGAATCAAGGATATTGACAGATACCATGAAAAGCT  
AGCAGTATCTCTGTATAGCTCTTCTCTTCCACCACACCTATCTGGTATTAAATTAGGTGT  
AATTAGCTCTGAAGAGCTTCCTTTATACTACAGGAGAAGTGTGAGCTAAATAGCAATTC  
TTTGGAGTCCATGAAAATGTCATGCAGAGTTCCTTGGGATTTGTTAATCCAGGAGTAGT  
TGAATTCCTTCTAGAAAACTAGGAATAGATGAATCTAATCCACCATCTTTAATGAGAGG  
CCTACAAAGCAAAAATTTCTCTGAAGATGATTTTGCTTCTGCATTGAGAAAAAACAATC  
CGCATCTTGGGCTTTAAATGTGTGAAGATCGGAGTTGACTATTTTAAAGTTGGACGCCA  
TGTGGATGCTATGAATGAATACAATAAAGCTTTGGAAATAGACAAACAAAACGTGGAAGC  
TTTGGTAGCTCGTGGAGCATTATATGCGACAAAAGGAAGTTTGAACAAAGCAATAGAAGA  
TTTTGAGCTTGCATTAGAAAATGTCCAATCACAGAAATGCAAGAAAATACCTCTGCCA  
GACACTTGTAGAGAGAGGAGGACAGTTAGAAGAAGAAGAAAAGTTTTAAATGCTGAAAG  
TACTATAAGAAAGCCTTGGCTTTGGATGAGACTTTTAAAGATGCAGAGGATGCTTTGCA  
GAACTTCATAAATATATGCAGAAATCTTTGGAATTAAGAGAAAAACAAGCTGAAAAGGA  
AGAAAAGCAGAAAACAAAGAAAATAGAAACAAGTGCAGAAAAGTTGCGTAAGCTCTTAA  
AGAAGAGAAGAGGCTAAAGAAGAAAAGAAGAAAATCAACTTCTTCTCAAGTGTTCCTC  
TGCTGATGAATCAGTGTCTTCATCATCATCCTCTTCTCTTCTGGTCACAAAAGGCATAA  
GAAACATAAGAGGAACCGTTCAGAGTCTTCTCGCAGTTCAGAAAGGCATTTCATCTAGGGC  
ATCCTCAAATCAGATAGATCAGAATAGGAAAGATGAGTGCTACCCAGTTCCAGCTAATAC  
TTCAGCATCTTTTCTTAACCATAAACAAGAAGTGGAGAACTACTGGGGAAGCAGGATAG  
GTTACAGTATGAAAAGACACAGATAAAAGAGAAAGATAGATGCCCTCTCTCTCATCTTC  
ACTTGAAATACCGGATGATTTTGGAGTGTACTCCTATTTATTTAAAAAGTTAACTATAAA  
ACAGCCTCAGGCAGGTCCTTCAGGAGATATTCCAGAAGAGGGCATTGTTATCATAGATGA  
CAGCTCCATTCATGTTACTGACCCTGAAGACCTTCAAGTGGGACAAGATATGGAGGTGGA  
AGACAGTGGTATTGATGATCCTGACCACGGG**TAG**GCTTAGGTTTATGTGTGTGTATGTGT  
CTTAGTTTTTAAACAAAAAATTAATAAAGTAAAAAACTAAAAATAGAAAAATGCTTAGAG  
AATAAGGATATAAAGAATATTTTGTGCAGTTGAACAATGAGTGCTTAAGCTAAATGTCA  
TCACAAAAGAGTAAAAAAATTTTACAAAATTAATAAAGTTTTAAAGTTAAAAAGCTCTAGG  
AAGCTAAGGTCAATTTATTATTGGAGAAATAAATTATTTTTATGAATTTACTGT



406/615

**FIGURE 402**

MDRDLLRQSLNCHGSSLLSLRSEQQDNPHFRSLLGSAEPARGPPPQHPLQGRKEKRVD  
NIEIQKFISKKADLLFALSWKSDAPATSEINEDSEDHYAIMPPLEQFMEIPSMRRELEFF  
RDIERGDIVIGRISSIREFGFFMVLICLGSGIMRDIAHLEITALCPLRDVPSHSHNGDPL  
SYYQTGDIIRAGIKDIDRYHEKLAVSLYSSSLPPHLSGIKLGVISSEELPLYRRSVELN  
SNSLESYENVMQSSLGFVNPGVVEFLLEKLGIDESNPPSLMRGLQSKNFSEDDFASALRK  
KQSASWALKCVKIGVDYFKVGRHVDAMNEYNKALEIDKQNEALVARGALYATKGS LNKA  
IEDFELALENCPTHNRNARKYLCQTLVERGGQLEEEEEKFLNAESYKKALALDETFKDAED  
ALQKLHKYMQKSLELREKQAEKEEKQKTKKIETSAEKLRLKLLKEEKRLKKRRKSTSSSS  
VSSADESVSSSSSSSSSGHKRHKHKRNRSESSRRHSSRASSNQIDQNRKDECYPVP  
ANTSASFNLHKQEVKLLGKQDRLOYEKTOIKEKDRCP LSSSSLEIPDDFGVSYL FKKL  
TIKQPQAGPSGDIPEEGIVIIDSSIHVTDPEDLQVGQDMEVEDSGIDDPDHG

**Important features of the protein:****Signal peptide:**

Amino acids 1-23

**Transmembrane domain:**

Amino acids 138-155

**N-glycosylation sites:**

Amino acids 288-292;508-512;542-546

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 300-304;472-476;473-477;517-521;598-602

**N-myristoylation sites:**

Amino acids 218-224;222-228;271-277;348-354

**Amidation site:**

Amino acids 52-56

**Cell attachment sequence:**

Amino acids 125-128

407/615

**FIGURE 403**

CCGAGGCGGGAGGAGCCCGAGGGGGCGCGAGCCCCGCATGAATCATTGTAGTCAATCATTTTC  
CAGTTCTCAGCCGCTCAGTTGTGATCAAGGGACACGTGGTTTCCGAAGTCCAGCTCAGAATA  
GGAAAATAACTTGGGATTTTATATTGGAAGAC**ATGG**ATCTTGCTGCCAACGAGATCAGCATTT  
ATGACAAACTTTTCAAGAGACTGTTGATTTGGTGAGACAGACCGGCCATCAGTGTGGCATGTCAG  
AGAAGGCAATTGAAAAATTTATCAGACAGCTGCTGGAAAAGAATGAACCTCAGAGACCCCCC  
CGCAGTATCCTCTCCTTATAGTTGTGTATAAGGTTCTCGCAACCTTGGGATTAATCTTGCTCA  
CTGCCTACTTTGTGATTCAACCTTTCAGCCCATTAGCACCTGAGCCAGTGCTTTCTGGAGCTC  
ACACCTGGCGCTCACTCATCCATCACATTAGGCTGATGTCCTTGCCCATTGCCAAGAAGTACA  
TGTCAGAAAATAAGGGAGTTCTCTGCATGGGGGTGATGAAGACAGACCCTTTCCAGACTTTG  
ACCCCTGGTGGACAAACGACTGTGAGCAGAAAGAGTCAGAGCCCATTCCTGCCAACTGCACTG  
GCTGTGCCCAGAAACACCTGAAGGTGATGCTCCTGGAAGACGCCCCAAGGAAATTTGAGAGGC  
TCCATCCACTGGTGTATCAAGACGGGAAAGCCCCTGTTGGAGGAAGAGATTCAGCATTTTTTGT  
GCCAGTACCCTGAGGCGACAGAAGGCTTCTCTGAAGGGTTTTTCGCAAGTGGTGGCGCTGCT  
TTCCTGAGCGGTGGTTCCCATTTCTTATCCATGGAGGAGACCTCTGAACAGATCACAAATGT  
TACGTGAGCTTTTTCTGTCTTCACTCACCTGCCATTTCCAAAAGATGCCTCTTTAAACAAGT  
GCTCCTTTCTTACCCAGAACCTGTTGTGGGGAGTAAGATGCATAAGATGCCTGACCTATTTA  
TCATTGGCAGCGGTGAGGCCATGTTGCAGCTCATCCCTCCCTTCCAGTGCCGAAGACATTGTC  
AGTCTGTGGCCATGCCAATAGAGCCAGGGGATATCGGCTATGTGACACCAACCCACTGGAAGG  
TCTACGTTATAGCCAGAGGGGTCCAGCCTTTGGTCATCTGCGATGGAACCGCTTTCTCAGAAC  
TG**TAG**GAAATAGAACTGTGCACAGGAACAGCTTCAGAGCCGAAAACCAGGTTGAAAGGGGAA  
AAATAAAAAACAAAAACGATGAAACTGCAAAAA

408/615

**FIGURE 404**

MDLAANEISIIYDKLSETVDLVRQTGHQCGMSEKAIEKFIRQLLEKNEPQRPPQYPLLIVVYK  
VLATLGLILLTAYFVIQPFSPLAPEPVLSGAHTWRSLIHHIRLMSLPIAKKYMSENKGVPLHG  
GDEDRPFPDFDPWWTNDCEQNESEPIPANCTGCAQKHLKVMILLEDAPRKFERLHPLVIKTGKP  
LLEEEIQHFLLCQYPEATEGFSEGFFAKWWRCFPERWFPPYPWRRPLNRSQMLRELFVPV'THL  
PFPKDASLNKCSFLHPEPVVGSKMHKMPDLFIIGSGEAMLQLIPPFQRRHCQSVAMPIEPGD  
IGYVDTTHWKVYVIARGVQPLVICDGTAFSEL

409/615

**FIGURE 405**

TGCCGGGCTGCGGGGCGCCTTGACTCTCCCTCCACCCTGCCTCCTCGGGCTCCACTCGTCTGCCCCCTGGACTCCC  
GTCTCCTCCTGTCTCCTCCGGCTTCCCAGAGCTCCCTCCTTATGGCAGCAGCTTCCCGCGTCTCCGGCGCAGCTTCT  
CAGCGGACGACCCCTCTCGCTCCGGGGCTGAGCCCAGTCCCTGGATGTTGCTGAAACTCTCGAGATCATGCGCGGG  
TTTGGCTGCTGCTTCCCCGCCGGGTGCCACTGCCACCGCCGCCCTCTGCTGCCGCCGTCCGCGGGATGCTCAG  
TAGCCCGCTGCCCGGCCCCCGGATCCTGTGTTCTCGGAAGCCGTTTGCTGCTGCAGAGTTGCACGAACTAGTC  
ATGGCTGCTGTTGGGAGTCCCCGCGGCAGTGCAGCAGCTGGACACTTTGCGAGGGCTTTTGCTGGCTGCTGCTGCTG  
CCCGTCATGCTACTCATCGTAGCCCGCCCGGTGAAGCTCGCTGCTTTCCTTACCTCCTTAAGTGACTGCCAAACG  
CCCACCGCTGGAATTGCTCTGTTATGATGACAGAGAAAATGATCTCTTCTCTGTGACACCAACACCTGTAAA  
TTTGATGGGGAATGTTTAAGAATTGGAGACACTGTGACTTGCCTCTGTTCAGTTCAAGTGCAACAATGACTATGTG  
CCTGTGTGTGGCTCCAATGGGGAGAGCTACCAGAATGAGTGTTACCTGCGACAGGCTGCATGCAAACAGCAGAGT  
GAGATACTTGTGGTGTGAGAAGGATCATGTGCCACAGATGCAGGATCAGGATCTGGAGATGGAGTCCATGAAGGC  
TCTGGAGAACTAGTCAAAAGGAGACATCCACCTGTGATATTTGCCAGTTTGGTGCAAGATGTGACGAAGATGCC  
GAGGATGTCTGGTGTGTGTGTAATATTGACTGTTCTCAAACCAACTTCAATCCCCTCTGCGCTTCTGATGGGAAA  
TCTTATGATAATGCATGCCAAATCAAAGAAGCATCGTGTGAGAAACAGGAGAAAATTGAAGTCATGCTTTGGGT  
CGATGTCAAGATAACACAACACTAACTACTAAGTCTGAAGATGGGCATTATGCAAGAACAGATTATGCAGAGAAT  
GCTAACAAATTAGAAGAAAGTGCCAGAGAACACCACATACCTTGTCCGGAACATTACAATGGCTTCTGCATGCAT  
GGGAAGTGTGAGCATTCTATCAATATGCAGGAGCCATCTTGCAAGGTGTGATGCTGGTTATACTGGACAACACTGT  
GAAAAAAGGACTACAGTGTTCTATACGTTGTTCCCGGTCCTGTACGATTTTCAAGTATGTCTTAATCGCAGCTGTG  
ATTGGAACAATTGAGATTGCTGTGCTGCTGTGTTGCTCCTCTGCATCACAAGGAAATGCCCCAGAAGCAACAGA  
ATTCACAGACAGAAGCAAAATACAGGGCACTACAGTTCAGACAATACAACAAGAGCGTCCACGAGGTTAATCTAA  
AGGGAGCATGTTTACAGTGGCTGGACTACCGAGAGCTTGGACTACACAATACAGTATTATAGACAAAAGAATAA  
GACAAGAGATCTACACATGTTGCCTTGCATTTGTGGTAATCTACACCAATGAAAACATGTACTACAGCTATATTT  
GATTATGTATGGATATATTTGAAATAGTATACATTGTCTTGATGTTTTTCTGTAATGTAAATAAACTATTTATA  
TCACACAATATAGTTTTTTCTTTCCCATGTATTTGTTATATATAATAAATACTCAGTGATGAG

410/615

**FIGURE 406**

MVLWESPRQCSSWTLCEGFCWLLLLPVMLLIVARPVKLAAPPTSLSDCQTPTGWNC SGY  
DDRENDLFLCDTNTCKFDGECLRIGDTVTVCVCQFKCNNDYVPVCGSNGESYQNECYLRQ  
AACKQQSEILVVSEGSCATDAGSGSGDGVHEGSGETSQKETSTCDICQFGAECDEDAED  
VWCVCNIDCSQTNFNPLCASDGKSYDNACQIKEASCQKQEKIEVMSLGRCQDNTTTTTK  
SEDGHYARTDYAENANKLEESAREHHIPCPEHYNGFCMHGKCEHSINMQEPSCRC DAGY  
TGQHCEKKDYSVLYVVPGPVRFQYVLIAAVIGTIQIAVICVVVLCITRKCPRSNRIHRQ  
KQNTGHYSSDNTTRASTRLI

411/615

**FIGURE 407**

CTCGCAGCCGAGCGCGGCCGGGGAAGGGCTCTCCTTCCAGCGCCGAGCACTGGGCCCTGGCAG  
ACGCCCCAAGATTGTTGTGAGGAGTCTAGCCAGTTGGTGAGCGCTGTAATCTGAACCAGCTGT  
GTCCAGACTGAGGCCCATTTGCATTGTTTAACATACTTAGAAAATGAAGTGTTCATTTTTAA  
CATTCTCTCCAATTGGTTTAATGCTGAATTACTGAAGAGGGCTAAGCAAACCAGGTGCTT  
GCGCTGAGGGCTCTGCAGTGGCTGGGAGGACCCCGGCGCTCTCCCCGTGTCTCTCCACGACT  
CGCTCGGCCCTCTGGAATAAAACACCCGCGAGCCCCGAGGGCCCAGAGGAGGCCGACGTGCC  
CGAGCTCTCCGGGGTCCCGCCCGCGAGCTTTCTTCTCGCCTTCGCATCTCCTCCTCGCGCG  
TCTTGAGACATGCCAGGAATAAAAAGGATACTACTGTTACCATTCTGGCTCTCTGTCTTCCAA  
GCCCTGGGAATGCACAGGCACAGTGCACGAATGGCTTTGACCTGGATCGCCAGTCAGGACAGT  
GTTTAGATATTGATGAATGCCGAACCATCCCCGAGGCCTGCCGAGGAGACATGATGTGTGTTA  
ACCAAATGGCGGGTATTTATGCATTCCCCGACAAACCCTGTGTATCGAGGGCCCTACTCGA  
ACCCCTACTCGACCCCTACTCAGGTCCGTACCCAGCAGCTGCCCCACCACTCTCAGCTCCAA  
ACTATCCCACGATCTCCAGGCCTCTTATATGCCGCTTTGGATACCAGATGGATGAAAGCAACC  
AATGTGTGGATGTGGACGAGTGTGCAACAGATTCCACACAGTGCAACCCCCACCCAGATCTGCA  
TCAATACTGAAGGCGGGTACACCTGCTCCTGCACCGACGGATATTGGCTTCTGGAAGGCCAGT  
GCTTAGACATTGATGAATGTGCTATGGTTACTGCCAGCAGCTCTGTGCGAATGTTTCTGGAT  
CCTATTCTTGTACATGCAACCCTGGTTTTTACCCTCAATGAGGATGGAAGGTCTTGCCAAGATG  
TGAACGAGTGTGCCACCGAGAACCCTGCGTGCAACCTGCGTCAACACCTACGGCTCTCTCA  
TCTGCCGCTGTGACCCAGGATATGAACCTGAGGAAGATGGCGTTCATTGCAGTGATATGGACG  
AGTGACAGCTTCTCTGAGTTCCTCTGCCAACATGAGTGTGTGAACCAGCCCGGCACATACTTCT  
GCTCCTGCCCTCCAGGCTACATCCTGCTGGATGACAACCGAAGCTGCCAAGACATCAACGAAT  
GTGAGCACAGGAACCACACGTGCAACCTGCAGCAGACGTGCTACAATTTACAAGGGGGCTTCA  
AATGCATCGACCCCATCCGCTGTGAGGAGCCTTATCTGAGGATCAGTGATAACCGCTGTATGT  
GTCCTGCTGAGAACCCTGGCTGCAGAGACCAGCCCTTTACCATCTTGTACCGGGACATGGACG  
TGGTGTGAGGACGCTCCGTTCCCGCTGACATCTTCAAATGCAAGCCACGACCCGCTACCCCTG  
GGGCCTATTACATTTTCCAGATCAAATCTGGGAATGAGGGCAGAGAATTTTACATGCGGCAAA  
CGGGCCCCATCAGTGCCACCCCTGGTGATGACACGCCCCATCAAAGGGCCCCGGGAAATCCAGC  
TGGACTTGGAATGATCACTGTCAACACTGTCATCAACTTCAGAGGCAGCTCCGTGATCCGAC  
TGCGGATATATGTGTGCGCAGTACCCATTCTGAGCCTCGGGCTGGAGCCTCCGACGCTGCCTCT  
CATTGGCACCAAGGGACAGGAGAAGAGAGGAAATAACAGAGAGAATGAGAGCGACACAGACGT  
TAGGCATTTCTCTGCTGAACGTTTCCCGAAGAGTCAGCCCCGACTTCCTGACTCTCACCTGTA  
CTATTGCAGACCTGTCACCCTGCAGGACTTGCCACCCCAAGTTCCTATGACACAGTTATCAA  
AAGTATTATCATTGCTCCCTGATAGAAGATTGTTGGTGAATTTTCAAGGCCTTCAGTTTATT  
TCCACTATTTTCAAAGAAAATAGATTAGGTTTGCGGGGGTCTGAGTCTATGTTCAAAGACTGT  
GAACAGCTTGCTGTCACTTCTTACCTCTTCCACTCCTTCTCTCACTGTGTTACTGCTTTGCA  
AAGACCCGGGAGCTGGCGGGGAACCCTGGGAGTAGCTAGTTTGCTTTTTCGCTACACAGAGAA  
GGCTATGTAAACAAACCACAGCAGGATCGAAGGGTTTTTAGAGAATGTGTTTCAAACCATGC  
CTGGTATTTTCAACCATAAAAGAAGTTTCAGTTGTCCTTAAATTTGTATAACGGTTTAATTCT  
GTCTTGTTCAATTTGAGTATTTTAAAAAATATGTCGTAGAATTCCTTCGAAAGGCCTTCAGA  
CACATGCTATGTTCTGTCTTCCCAAACCCAGTCTCCTCTCCATTTTAGCCAGTGTTTCTTT  
GAGGACCCCTTAATCTTGCTTCTTTAGAATTTTACCCTTGGATTGGAATGCAGAGGTCT  
CCAAACTGATTAAATATTTGAAGAGA

412/615

**FIGURE 408**

MPGIKRILTVTILALCLPSPGNAQAQCTNGFDLDRQSGQCLDIDECRTIPEACRGDMMCVNQNGG  
GYLCIPRTNPVYRGPYSNPYSTPYSGPYPAAAPPLSAPNYPTISRPLICRFGYQMDESNQCV  
DVDECATDSHQCNPTQICINTEGGYTCSTDGWLLLEGQCLDIDECRYGYCQQLCANVPGSYS  
CTCNPGFTLNEDGRSCQDVNECATENPCVQTCVNTYGSLICRCDPGYELEEDGVHCSDMDECS  
FSEFLCQHECVNQPGTYFCSCPPGYILLDDNRSCQDINECEHRNHTCNLQQTTCYNLQGGFKCI  
DPIRCEEPYLRISDNRCMCPAENPGCRDQPTILYRDMDVVSGRSVPADIFQMQATTRYPGAY  
YIFQIKSGNEGREFYMRQTGPISATLVMTRPIKGPREIQLDLEMITVNTVINFRGSSVIRLRI  
YVSQYPF

**Important features of the protein:****Signal peptide:**

amino acids 1-25

**N-glycosylation sites.**

amino acids 283-287, 296-300

**N-myristoylation sites.**amino acids 21-27, 64-70, 149-155, 186-192, 226-232, 242-248,  
267-273, 310-316**Aspartic acid and asparagine hydroxylation sites.**

amino acids 144-156, 181-193, 262-274

**Cell attachment sequence.**

amino acids 54-57

**Calcium-binding EGF-like.**

amino acids 131-166, 172-205, 211-245, 251-286

413/615

**FIGURE 409**

CCCACGCGTCCGCGGACGCGTGGGTGCGACTAGTTCTAGATCGCGAGCGGCCGCCGCGGCTCA  
GGGAGGAGCACCGACTGCGCCGCACCCCTGAGAGATGGTTGGTGCCATGTGGAAGGTGATTGTT  
TCGCTGGTCCTGTTGATGCCTGGCCCTGTGATGGGCTGTTTCGCTCCCTATACAGAAGTGTT  
TCCATGCCACCTAAGGGAGACTCAGGACAGCCATTATTTCTCACCCCTTACATTGAAGCTGGG  
AAGATCCAAAAAGGAAGAGAATTGAGTTTGGTCGGCCCTTTCCCAGGACTGAACATGAAGAGT  
TATGCCGGCTTCCTCACCGTGAATAAGACTTACAACAGCAACCTCTTCTTCTGGTTCTTCCCA  
GCTCAGATACAGCCAGAAGATGCCCCAGTAGTTCTCTGGCTACAGGGTGGGCCGGGAGGTTCA  
TCCATGTTTGGACTCTTTGTGGAACATGGGCCTTATGTTGTCACAAGTAACATGACCTTGCGT  
GACAGAGACTTCCCCTGGACCACAACGCTCTCCATGCTTTACATTGACAATCCAGTGGGCACA  
GGCTTCAGTTTACTGATGATACCCACGGATATGCAGTCAATGAGGACGATGTAGCACGGGAT  
TTATACAGTGCATAATTGAGTTTTCAGATATTTCTGAATATAAAAAATAATGACTTTTAT  
GTCAGTGGGGAGTCTTATGCAGGGAAATATGTGCCAGCCATTGCACACCTCATCCATTCCCTC  
AACCCTGTGAGAGAGGTGAAGATCAACCTGAACGGAATTGCTATTGGAGATGGATATTCTGAT  
CCCGAATCAATTATAGGGGGCTATGCAGAATTCCTGTACCAAATTGGCTTGTTGGATGAGAAG  
CAAAAAAGTACTTCCAGAAGCAGTGCCATGAATGCATAGAACACATCAGGAAGCAGAACTGG  
TTTGAGGCCTTTGAAATACTGGATAAACTACTAGATGGCGACTTAACAAGTGATCCTTCTTAC  
TTCCAGAATGTTACAGGATGTAGTAATTACTATAACTTTTTGCGGTGCACGGAACCTGAGGAT  
CAGCTTTACTATGTGAAATTTTTGTCACTCCCAGAGGTGAGACAAGCCATCCACGTGGGGAAT  
CAGACTTTTAATGATGGAACATAGTTGAAAAGTACTTGCGAGAAGATACAGTACAGTCAGTT  
AAGCCATGGTTAACTGAAATCATGAATAATTATAAGGTTCTGATCTACAATGGCCAACTGGAC  
ATCATCGTGGCAGCTGCCCTGACAGAGCGCTCCTTGATGGGCATGGACTGGAAAGGATCCAG  
GAATACAAGAAGGCAGAAAAAAAGTTTGGAAGATCTTTAAATCTGACAGTGAAGTGGCTGGT  
TACATCCGGCAAGCGGGTGACTCCATCAGGTAATTATTCGAGGTGGAGGACATATTTTACCC  
TATGACCAGCCTCTGAGAGCTTTTGACATGATTAATCGATTCAATTTATGGAAAAGGATGGGAT  
CCTTATGTTGGATAAACTACCTTCCCAAAAGAGAACATCAGAGGTTTTTCATTGCTGAAAAGAA  
AATCGTAAAAACAGAAAATGTCATAGGAATAAAAAAATTATCTTTTCATATCTGCAAGATTTT  
TTTCATCAATAAAAAATTATCCTTGAAACAAGTGAGCTTTTGTTTTGGGGGGAGATGTTTACT  
ACAAAATTAACATGAGTACATGAGTAAGAATTACATTATTTAACTTAAAGGATGAAAGGTATG  
GATGATGTGACACTGAGACAAGATGTATAAATGAAATTTTAGGGTCTTGAATAGGAAGTTTTA  
ATTTCTTCTAAGAGTAAGTGAAAAGTGCAGTTGTAACAAACAAAGCTGTAACATCTTTTCTG  
CCAATAACAGAAGTTTGGCATGCCGTGAAGGTGTTTGGAATATTATTGGATAAGAATAGCTC  
AATTATCCCAAATAAATGGATGAAGCTATAATAGTTTGGGGAAAAGATTCTCAAATGTATAA  
AGTCTTAGAACAAAAGAATTCTTTGAAATAAAAAATATTATATATAAAAGTAAAAA



414/615

**FIGURE 410**

MVGAMWKVIVSLVLLMPGPCDGLFRSLYRSVSMPPKGDGSGQPLFLTPYIEAGKIQKGRELSLV  
GPFPGLNMKSYAGFLTIVNKTYSNLFWFPPAQIQPEDAPVVLWLQGGPGGSSMFGLFVEHGP  
YVVTSNMTLRDRDFPWTTLTMLYIDNPVGTGFSFTDDTHGYAVNEDDVARDLYSALIQQFFQI  
FPEYKNNDFYVTGESYAGKYVPATAHLIHSNPNVREVKINLNGIAIGDGYSDPESIIGGYAEF  
LYQIGLLDEKQKKYFQKQCHECIEHIRKQNWFEAFEILDKLLDGDLTSDPSYFQNVGTGCSNYY  
NFLRCTEPEDQLYYVKFLSLPEVRQAIHVGNOTFNDGTIVEKYLREDTVQSVKPWLTEIMNNY  
KVLIYNGQLDIIVAAALTERSLMGMDWKGSQYKKAEEKVWKIFKSDSEVAGYIRQAGDFHQV  
IIRGGGHILPYDQPLRAFDMINRFIYGKGWDPYVG

**Signal sequence:**

amino acids 1-22

**N-glycosylation site.**

amino acids 81-85, 132-136, 307-311, 346-350

**Casein kinase II phosphorylation site.**amino acids 134-138, 160-164, 240-244, 321-325, 334-338, 348-352,  
353-357, 424-428**Tyrosine kinase phosphorylation site.**

amino acids 423-432

**N-myristoylation site.**

amino acids 22-28, 110-116, 156-162, 232-238

**Serine carboxypeptidases, serine active site.**

amino acids 200-208

**Crystallins beta and gamma 'Greek key' motif signature.**

amino acids 375-391

415/615

**FIGURE 411**

GCAAGCCAAGGCGCTGTTTGAGAAGGTGAAGAAGTTCCGGACCCATGTGGAGGAGGGGGACATTGTGTACCGCCT  
CTACATGCGGCAGACCATCATCAAGGTGATCAAGTTCATCCTCATCATCTGCTACACCGTCTACTACGTGCACAA  
CATCAAGTTCGACGTGGACTGCACCGTGGACATTGAGAGCCTGACGGGCTACCGCACCTACCGCTGTGCCACCC  
CCTGGCCACACTCTTCAAGATCCTGGCGTCCTTCTACATCAGCCTAGTCATCTTCTACGGCCTCATCTGCATGTA  
CACACTGTGGTGGATGCTACGGCGCTCCCTCAAGAAGTACTCGTTTGAGTCGATCCGTGAGGAGAGCAGCTACAG  
CGACATCCCCGACGTCAAGAACGACTTCGCCTTCATGCTGCACCTCATTGACCAATACGACCCGCTCTACTCCAA  
GCGCTTCGCGCTCTTCTGTGCGGAGGTGAGTGAGAACAAGCTGCGGCAGCTGAACCTCAACAACGAGTGGACGCT  
GGACAAGCTCCGGCAGCGCTCACCAAGAACGCGCAGGACAAGCTGGAGCTGCACCTGTTTCATGCTCAGTGGCAT  
CCCTGACACTGTGTTTGACCTGGTGGAGCTGGAGGTCTCAAGCTGGAGCTGATCCCCGACGTGACCATCCCCGCC  
CAGCATTGCCAGCTCACGGGCTCAAGGAGCTGTGGCTCTACCACACAGCGGCCAAGATTGAAGCGCTGCGCT  
GGCCTTCCTGCGCGAGAACCTGCGGGCGCTGCACATCAAGTTCACCGACATCAAGGAGATCCCGCTGTGGATCTA  
TAGCCTGAAGACACTGGAGGAGCTGCACCTGACGGGCAACCTGAGCGCGGAGAACAACCGCTACATCGTCATCGA  
CGGGCTGCGGGAGCTCAAACGCTCAAGGTGCTGCGGCTCAAGAGCAACCTAAGCAAGCTGCCACAGGTGGTCAC  
AGATGTGGGCGTGCACCTGCAGAAGCTGTCCATCAACAATGAGGGCAGCAAGCTCATCGTCTCAACAGCCTCAA  
GAAGATGGCGAACCTGACTGAGCTGGAGCTGATCCGCTGCGACCTGGAGCGCATCCCCACTCCATCTTCAGCCT  
CCACAACCTGCAGGAGATTGACCTCAAGGACAACAACCTCAAGACCATCGAGGAGATCATCAGCTTCAGCACCT  
GCACCGCTCACCTGCCTTAAGCTGTGGTACAACCACATCGCCTACATCCCCATCCAGATCGGCAACCTCACCAA  
CCTGGAGCGCTCTACCTGAACCGCAACAAGATCGAGAAGATCCCCACCCAGCTCTTCTACTGCCGAAGCTGCG  
CTACCTGGACCTCAGCCACAACAACCTGACCTTCCTCCCTGCCGACATCGGCCTCCTGCAGAACCTCCAGAACCT  
AGCCATCACGGCAACCGGATCGAGACGCTCCCTCCGGAGCTCTTCCAGTGCCGGAAGCTGCGGGCCCTGCACCT  
GGGCAACAACGTGCTGCAGTCACTGCCCTCCAGGGTGGGCGAGCTGACCAACCTGACGCAGATCGAGCTGCGGGG  
CAACCGGCTGGAGTGCTGCCCTGTGGAGCTGGGCGAGTGCCCACTGCTCAAGCGCAGCGGCTTGGTGGTGGAGGA  
GGACCGCTTCAACACACTGCCACCGAGGTGAAGGAGCGGCTGTGGAGGCTGACAAGGAGCAGGCTGAGCGGAG  
CGCGGCCAGCACAGCAAGCAGCAGGACCGCTGCCAGTCTCAGGCCCGGAGGGGAGGCTAGCTTCTCCCAG  
AACTCCCGGACAGCCAGGACAGCCTCGCGGCTGGGCGAGGCTGGGCGCGCTTGTGAGTCAGGCCAGAGCGAGA  
GGACAGTATCTGTGGGGCTGGCCCTTTTCTCCCTCTGAGACTCACGTCCCCAGGGCAAGTGCTTGTGGAGGAG  
AGCAAGTCTCAAGAGCGCAGTATTTGGATAATCAGGGTCTCCTCCCTGGAGGCCAGCTCTGCCCCAGGGGCTGAG  
CTGCCACCAGAGGTCTGGGACCTCACTTTAGTTCTTGGTATTTATTTTCTCCATCTCCCACCTCCTTCATCC  
AGATACTTATACATTCCCAAGAAAGTTCAAGCCAGATGGAAGGTGTTCAAGGAAAGGTGGGCTGCCTTTTCCCC  
TTGTCTTATTTAGCGATGCCGCCGGGCATTTAACACCCACCTGGACTTCAGCAGAGTGGTCCGGGGCGAACCCAG  
CCATGGGACGGTCACCCAGCAGTGCCGGCTGGGCTCTGCGGTGCGGTCCACGGGAGAGCAGGCTCCAGCTGGA  
AAGGCCAGGCTGGAGCTTGCTCTTCAGTTTTTGTGGCAGTTTTAGTTTTTGTTTTTTTTTTTTTTAAATCAA  
AAACAATTTTTTTTTTAAAAAAGCTTTGAAATGGATGGTTTGGGTATTAAGGAAAAAAGCTTAAAAA  
AAAAGACACTAACGGCCAGTGAGTTGGAGTCTCAGGGCAGGGTGGCAGTTTCCCTTGAGCAAAGCAGCCAGAGT  
TGAAGTGTGTTTCTTCCCTGGGCGCAGGGTGAGGGTGTCTTCCGATCTGGTGTGACCTTGGTCCAGGAGTT  
CTATTTGTTTCTTGGGAGGGAGGTTTTTGTGTTGTTTTTGGGTTTTTGGTGTCTTGTCTTCTCTCTCC  
ATGTGCTTGGCAGGCACTCATTTCTGTGGCTGTGCGCCAGAGGGAATGTTCTGGAGCTGCCAAGGAGGGAGGAG  
ACTCGGGTTGGCTAATCCCCGGATGAACGGTGCTCCATTGCGACCTCCCTCCTCGTGCCTGCCCTGCCCTCTCCA  
CGCACAGTGTTAAGGAGCCAAGAGGAGCCACTTCGCCAGACTTTGTTTCCCCACCTCCTGCGGCATGGGTGTGT  
CCAGTGCCACCGCTGGCCTCCGCTGCTTCCATCAGCCCTGTGCGCACCTGGTCTTTCATGAAGAGCAGACCTTA  
GAGGCTGGTGGGAATGGGGAGGTGCGCCCTGGGAGGGCAGGCGTTGGTTCCAAGCCGTTCCCGTCCCTGGCGC  
CTGGAGTGACACAGCCAGTCCGACCTGGTGGCTGGAAGCCAACCTGCTTTAGATCACTCGGTCCCCACCTT  
AGAAGGGTCCCCGCTTAGATCAATCACGTGGACACTAAGGCACGTTTTAGAGTCTTGTCTTAATGATTATGT  
CCATCCGTCTGTCCGTCCATTTGTGTTTTCTGCGTGTGTATTGGATATAATCCTCAGAAATAATGCACACTAG  
CCTCTGACAACCATGAAGCAAAATCCGTTACATGTGGGTCTGAACTGTAGACTCGGTACAGTATCAAATAAA  
ATCTATAACAGAAAAA

416/615

**FIGURE 412**

MRQTIKVIKFILIIICYTVYYVHNIKFDVDCTVDIESLTGYRTYRCAHPLATLFKILASFYIS  
LVIFYGLICMYTLWWMLRRSLKKYSFESIREESSYSDIPDVKNDFAFMLHLIDQYDPLYSKRF  
AVFLSEVSENKLRQLNLNNEWTLDKLRQRLTKNAQDKLELHLFMLSGIPDPTVFDLVELEVLKL  
ELIPDVTIPPSIAQLTGLKELWLYHTAAKIEAPALAFLENLRLALHIKFTDIKEIPLWIYSLK  
TLEELHLTGNLSEAENNRIVIDGLRELKRLKVLRLKSNLSKLPQVVTDVGVHLQKLSINNEG  
KLIVNSLKKMANLTELELIRCDLERIPHSIFSLHNLQEIDLKDNNLKTIEEIIISFOHLHRLT  
CLKLWYNHIAIYIPIQIGNLTNLERLYLNRNKIEKIPTQLFYCRKLRYLDLSHNNLTFLPADIG  
LLQNLQNLAITANRIETLPPELFQCRKLRLALHLGNNVLQSLPSRVGELTNLTQIELRGNRLEC  
LPVELGECPLLKRSGLVVEEDLFENTLPPEVKERLWRADKEQA

**Transmembrane domain:**

amino acids 51-75 (type II)

**N-glycosylation site.**

amino acids 262-266, 290-294, 328-332, 396-400, 432-436, 491-495

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 85-89

**Casein kinase II phosphorylation site.**amino acids 91-95, 97-101, 177-181, 253-257, 330-334, 364-368,  
398-402, 493-497**N-myristoylation site.**

amino acids 173-179, 261-267, 395-401, 441-447

417/615

**FIGURE 413**

GAATCATCCACGCACCTGCAGCTCTGCTGAGAGAGTGCAAGCCGTGGGGGTTTTGAGCTCATC  
TTCATCATTTCATATGAGGAAATAAGTGGTAAATCCTTGGAAATACAATGAGACTCATCAGAA  
ACATTTACATATTTTGTAGTATTGTTATGACAGCAGAGGGTGATGCTCCAGAGCTGCCAGAAG  
AAAGGGAACCTGATGACCAACTGCTCCAACATGTCTCTAAGAAAGGTTCCCGCAGACTTGACCC  
CAGCCACAACGACACTGGATTTATCCTATAACCTCCTTTTTCAACTCCAGAGTTCAGATTTTC  
ATTCTGTCTCCAAACTGAGAGTTTTGATTCTATGCCATAACAGAATTCACAGCTGGATCTCA  
AAACCTTTGAATTCACAAGGAGTTAAGATATTTAGATTTGTCTAATAACAGACTGAAGAGTG  
TAACTTGGTATTTACTGGCAGGTCTCAGGTATTTAGATCTTTCTTTAATGACTTTGACACCA  
TGCCTATCTGTGAGGAAGCTGGCAACATGTCACACCTGGAAATCCTAGGTTTGAGTGGGGCAA  
AAATACAAAATCAGATTTCCAGAAAATGTCTCATCTGATCTAATACTGTCTTCTTAGGAT  
TCAGAACTCTTCTCATTATGAAGAAGGTAGCCTGCCATCTTAAACACAACAAAACCTGCACA  
TTGTTTTACCAATGGACACAAATTTCTGGGTCTTTTTGCGTGATGGAATCAAGACTTCAAAAA  
TATTAGAAATGACAAATATAGATGGCAAAGCCAATTTGTAAGTTATGAAATGCAACGAAATC  
TTAGTTTAGAAAATGCTAAGACATCGGTTCTATTGCTTAATAAAGTTGATTTACTCTGGGACG  
ACCTTTTCTTATCTTACAATTTGTTTGGCATAACATCAGTGGAAACACTTTTCAGATCCGAAATG  
TGACTTTTGGTGGTAAGGCTTATCTTGACCACAAATTCATTTGACTACTCAAATACTGTAATGA  
GAACTATAAAATTTGGAGCATGTACATTTTCAGAGTGTTTTACATTCAACAGGATAAAATCTATT  
TGCTTTTGACCAAAATGGACATAGAAAACCTGACAATATCAAATGCACAAATGCCACACATGC  
TTTTCCCGAATTATCCTACGAAATTCCAATATTTAAATTTTGCCAATAATATCTTAACAGACG  
AGTTGTTTAAAGAAGTATCCAACCTGCCTCACTTGAAAACCTCTCATTTTGAATGGCAATAAAC  
TGGAGACACTTTCTTTAGTAAGTTGCTTTGCTAACAAACACACCCTTGGAAACACTTGGATCTGA  
GTCAAATCTATTACAACATAAAAATGATGAAAATTGCTCATGGCCAGAAACTGTGGTCAATA  
TGAATCTGTCAACAATAAATTGTCTGATTCTGTCTTCAGGTGCTTGCCCCAAAGTATTCAAA  
TACTTGACCTAAATAATAACCAATCCAACTGTACCTAAAGAGACTATTCATCTGATGGCCT  
TACGAGAACTAAATATTGCATTTAATTTTCTAACTGATCTCCCTGGATGAGTCAATTCAGTA  
GACTTTTCAGTTCTGAACATTGAAATGAACCTTCACTCTCAGCCCATCTCTGGATTTTGTTCAGA  
GCTGCCAGGAAGTTAAACTCTAAATGCGGGAAGAAATCCATTCCGGTGTACCTGTGAATTAA  
AAAATTTTCATTTCAGCTTGAACATATTCAGAGGTTCATGATGGTTGGATGGTCAGATTCATACA  
CCTGTGAATACCCTTTAAACCTAAGGGGAAGTGGTTAAAGACGTTTCATCTCCACGAATTAT  
CTTGCAACACAGCTCTGTTGATTGTCAACATTTGTGGTTATTATGCTAGTTCTGGGGTGGCTG  
TGGCCTTCTGCTGTCTCCACTTTGATCTGCCCTGGTATCTCAGGATGCTAGGTCAATGCACAC  
AAACATGGCACAGGGTTAGGAAAACAACCCAAGAACAACCTCAAGAGAAATGTCCGATTCACG  
CATTTATTTTCATACAGTGAACATGATTCTCTGTGGGTGAAGAATGAATTGATCCCCAATCTAG  
AGAAGGAAGATGGTTCTATCTTGATTTGCCCTTTATGAAAGCTACTTTGACCCTGGCAAAAGCA  
TTAGTGAAAATATTGTAAGCTTCATTGAGAAAAGCTATAAGTCCATCTTTGTTTTGTCTCCCA  
ACTTTGTCCAGAATGAGTGGTGCCATTATGAATTTCTACTTTGCCCCACACAATCTCTTCCATG  
AAAATTTCTGATCATATAATTCTTATCTTACTGGAACCCATTCCATTCTATTGCATTCCCACCA  
GGTATCATAACTGAAAGCTCTCCTGGAAAAAAGCATACTTGGAAATGGCCCAAGGATAGGC  
GTAAATGTGGGCTTTTCTGGGCAACCTTCGAGCTGCTATTAATGTTAATGTATTAGCCACCA  
GAGAAATGTATGAACCTGCAGACATTACAGAGTTAAATGAAGAGTCTCGAGGTTCTACAATCT  
CTCTGATGAGAACAGATTGTCTATAAAATCCCACAGTCCCTTGGGAAGTTGGGGACCACATACA  
CTGTTGGGATGTACATTGATACAACTTTATGATGGCAATTTGACAATATTTATTAAATAAA  
AAATGGTTATTCCCTTCATATCAGTTTCTAGAAGGATTTCTAAGAATGTATCCTATAGAAACA  
CCTTCACAAGTTTATAAGGGCTTATGGAAAAGGTGTTTCATCCAGGATTGTTTATAATCATG  
AAAAATGTGGCCAGGTGCAGTGGCTCACTCTTGTAATCCCAGCACTATGGGAGGCCAAGGTGG  
GTGACCCACGAGGTCAAGAGATGGAGACCATCCTGGCCAACATGGTGAAACCCCTGTCTCTACT  
AAAAATACAAAATTAGCTGGGCGTGATGGTGCACGCCTGTAGTCCCAGCTACTTGGGAGGCT  
GAGGCAGGAGAATCGCTTGAACCCGGGAGGTGGCAGTTGCAGTGAGCTGAGATCGAGCCACTG  
CACTCCAGCCTGGTGACAGAGCGAGACTCCATCTCAAAAAAAAAAGAAAAAAAAAAGAAAAA  
ATGGAAAACATCCTCATGGCCACAAAATAAGGTCTAATTCATAAATTATAGTACATTAATGT  
AATATAATATTACATGCCACTAAAAAGAATAAGGTAGCTGTATATTTCCCTGGTATGGAAAAA  
CATATTAATATGTTATAAACTATTAGGTGGTGCAAACTAATTGTGGTTTTTGGCATTGAAA  
TGGCATTGAAATAAAAGTGTAAAGAAATCTATACCAGATGTAGTAACAGTGGTTTTGGTCTGG  
GAGGTTGATTACAGGGAGCATTGATTCTATGTTGTGTATTCTATAATGTTTGAATTGTT  
TAGAATGAATCTGTATTTCTTTTATAAGTAGAAAAAATAAAGATAGTTTTTACAGCCT

418/615

**FIGURE 414**

MRLIRNIYIFCSIIVMTAEGDAPELPEERELMTNCSNMSLRKVPADLTPATTTLDLSYNLLFQL  
QSSDFHSVSKLRVLILCHNRIQQDLKTFFFNKELRYLDLSNNRLKSVTWYLLAGLRYLDLSF  
NDFDTMPICEEAGNMSHLEILGLSGAKIQKSDFQKIAHLHLNTVFLGFRTLPHYEEGSLPILN  
TTKLHIVLPMDTNFVLLRDGIKTSKILEMTNIDGKSQFVSYEMQRNLSLENAKTSVLLLNKV  
DLLWDDLFLILQFVWHTSVEHFQIRNVTFGGKAYLDHNSFDYSNTVMRTIKLEHVHFRVFYIQ  
QDKIYLLLTKMDIENLTISNAQMPHMLFPNYPTKFQYLNFANNILTDELFKRTIQLPHLKTLI  
LNGNKLETLSLVSCFANNTPLEHLDLSQNLQHKNDENC SWPETVVNMNLSYNKLSDSVFRCL  
PKSIQILDNLNNQIQTVPKETIHLMALRELNIAFNFLTDLPGCSHF SRLSVLNIEMNFILSPS  
LDFVQSCQEVKTLNAGRNPFRCTCELKNFIQLETYSEVMMVGWSDSYTCEYPLNLRGTRLKDV  
HLHELSCNTALLIVTIVVIMLVGLAVAFCCCLHFDLPWYLRMLGQCTQTWHRVRKTTQEQLKR  
NVRFHAFISYSEHDSLWVKNELIPNLEKEDGSILICLYESYFDPGKSISENIVSFIEKSYKSI  
FVLSPNFVQNEWCHYEFYFAHHNLFHENS DHIILILLEPIPFYCIPTRYHKLKALLEKKAYLE  
WPKDRRKCGLFWANLRAAINVNVLATREMYELQTFTELNEESRGSTISLMRTDCL

419/615

**FIGURE 415**

CGGACGCGTGGGCGGACGCGTGGGCCTGGGCAAGGGCCGGGGCGCCGGGCGGAGCCACCTCTTCCCCTCCCCCGC  
TTCCCTGTCGCGCTCCGCTGGCTGGACGCGCTGGAGGAGTGGAGCAGCACCCGGCCGGCCCTGGGGGCTGACAGT  
CGGCAAAGTTTGGCCCCGAAGAGGAAGTGGTCTCAAACCCCGGCAGGTGGCGACCAGGCCCAGACCAGGGGCGCTCG  
CTGCCTGCGGGCGGGCTGTAGGCGAGGGCGCGCCCCAGTGCCGAGACCCGGGGCTTCAGGAGCCCGCCCCGGGAG  
AGAAGAGTGGGCGGGCGGACGGAGAAAACTCAAAGTTGGCGAAAGGCACCGCCCCCTACTCCCAGGGCTGCCG  
CCGCCCTCCCCGCCCCAGCCCTGGCATCCAGAGTACGGGTGAGCCCGGGCCATGGAGCCCCCTGGGGAGGCGG  
CACCAGGGAGCCTGGGCGCCCCGGGGCTCCGCCGCGACCCCATCGGGTAGACCACAGAAGCTCCGGGACCCCTTCCG  
GCACCTCTGGACAGCCAGGATGCTGTTGGCCACCCTCCTCCTCCTCCTTGGAGGCGCTCTGGCCCATCCAG  
ACCGGATTATTTTCAAATCATGCTTGTGAGGACCCCCAGAGTGTCTTAGAAGTGCAGGGCACCTTACAGA  
GGCCCCCTGGTCCGGGACAGCCGCACCTCCCCTGCCAACTGCACCTGGCTCATCCTGGGCAGCAAGGAACAGACTG  
TCACCATCAGGTTCCAGAAGCTACACCTGGCCTGTGGCTCAGAGCGCTTAACCCCTACGCTCCCCTCTCCAGCCAC  
TGATCTCCCTGTGTGAGGCACCTCCCAGCCCTCTGCAGCTGCCCGGGGGCAACGTCACCATCACTTACAGCTATG  
CTGGGGCCAGAGCACCCCATGGGCCAGGGCTTCTGCTCTCTACAGCCAAGATTGGCTGATGTGCCTGCAGGAAG  
AGTTTCAGTGCCTGAACCACCGCTGTGTATCTGCTGTCCAGCGCTGTGATGGGGTTGATGCCCTGGCGATGGCT  
CTGATGAAGCAGGTTGCAGCTCAGACCCCTTCCCTGGCCTGACCCCAAGACCCGTCCTCCCTGCCTTGCAATG  
TCACCTTGGAGGACTTCTATGGGGTCTTCTCCTCTCCTGGATATACACACCTAGCCTCAGTCTCCACCCCCAGT  
CCTGCCATTGGCTGCTGGACCCCCATGATGGCCGGCGGCTGGCCGTGCGCTTACAGCCCTGGACTTGGGCTTTG  
GAGATGCAGTGCATGTGTATGACGGCCCTGGGCCCTTGAGAGCTCCCGACTACTGCGTAGTCTCACCCACTTCA  
GCAATGGCAAGGCTGTCACTGTGGAGACACTGTCTGGCCAGGCTGTTGTGCTCTACCACACAGTTGCTTGGAGCA  
ATGGTCGTGGCTTCAATGCCACCTACCATGTGCGGGGCTATTGCTTGCTTGGGACAGACCCTGTGGCTTAGGCT  
CTGGCCTGGGAGCTGGCGAAGGCCTAGGTGAGCGCTGTACAGTGAGGCACAGCGCTGTGACGGCTCATGGGACT  
GTGCTGACGGCACAGATGAGGAGGACTGCCCAGGCTGCCACCTGGACACTTCCCCTGTGGGGCTGCTGGCACCT  
CTGGTGCCACAGCCTGTACCTGCCTGCTGACCGCTGCAACTACCAGACTTTCTGTGCTGATGGAGCAGATGAGA  
GACGCTGTGCGCATTTGCCAGCCTGGCAATTTCCGATGCCGGGACGAGAAGTGCGTGTATGAGACGTGGGTGTGCG  
ATGGGCAGCCAGACTGTGCGGACGGCAGTGATGAGTGGGACTGCTCCTATGTTCTGCCCCGCAAGGTCATTACAG  
CTGCAGTCATTGGCAGCCTAGTGTGCGGCCTGCTCCTGGTCATCGCCCTGGGCTGCACCTGCAAGCTCTATGCCA  
TTTCGACCCAGGAGTACAGCATCTTTGCCCCCTCTCCCGATGGAGGCTGAGATTGTGCAGCAGCAGGCACCCC  
CTTCTACGGGCAGCTCATTTGCCAGGGTGCCATCCCACCTGTAGAAGACTTTCTACAGAGAATCCTAATGATA  
ACTCAGTGCTGGGCAACCTGCGTTCTCTGCTACAGATCTTACGCCAGGATATGACTCCAGGAGGTGGCCCCAGGTG  
CCCGCCGTGCTCAGCGGGGCGGCTTGATGCGACGCTGGTACGCCGTCTCCGCCGCTGGGGCTTGCTCCCTCGAA  
CCAAACCCCCGGCTCGGGCCTCTGAGGCCAGATCCCAGGTACACCTTCTGCTGCTCCCCTTGAGGCCCTAGATG  
GTGGCACAGGTCCAGCCCGTGAGGGCGGGGCGAGTGGGTGGGCAAGATGGGGAGCAGGCACCCCACTGCCCATCA  
AGGCTCCCCTCCCATCTGCTAGCACGTCTCCAGCCCCACTACTGTCCCTGAAGCCCCAGGGCCACTGCCCTCAC  
TGCCCCCTAGAGCCATCACTATTGTCTGGAGTGGTGCAGGCCCTGCGAGGCCGCTGTTGCCAGCCTGGGGCCCC  
CAGGACCAACCCGGAGCCCCCTGGACCCACACAGCAGTCTGGCCCTGGAAGATGAGGACGATGTGCTACTGG  
TGCCACTGGCTGAGCCGGGGGTGTGGTAGCTGAGGCAGAGGATGAGCCACTGCTTACCTGAGGGGACCTGGGGG  
CTCTACTGAGGCCCTCTCCCCTGGGGGCTCTACTCATAGTGGCACAACTTTTAGAGGTGGGTGAGCCTCCCCTCC  
ACCACTTCCCTTCCCTGTCCCTGGATTTCAGGGACTTGGTGGGCCTCCCGTTGACCCTATGTAGCTGCTATAAAGT  
TAAGTGTCCCTCAGGCAGGGAGAGGGCTCACAGAGTCTCCTCTGTACGTGGCCATGGCCAGACACCCCACTCCCT  
TCACCACCACCTGCTCCCCACGCCACCACATTTGGGTGGCTGTTTTAAAAAGTAAAGTTCTTAGAGGATCATA  
GGTCTGGACACTCCATCCTTGCCAAACCTTACCCAAAAGTGGCCTTAAGCACCGGAATGCCAATTAAC TAGAGA  
CCCTCCAGCCCCCAAGGGGAGGATTGGGCGAAGCTGAGGTTTTGCCATCCACAATCCCTCCTACAGGGCCTGG  
CTCACAAAAAGAGTGCAACAAATGCTTCTATTCCATAGCTACGGCATTGCTCAGTAAGTTGAGGTCAAAAATAAA  
GGAATCATACATCTC

420/615

**FIGURE 416**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA49631

&lt;subunit 1 of 1, 713 aa, 1 stop

&lt;MW: 76193, pI: 5.42, NX(S/T): 4

MLLATLLLLLLGGALAHDPDRIIFPNHACEDPPAVLLEVQGTLQRPLVRDSRTSPANCTWLILG  
SKEQTVTIRFQKLHLACGSERLTLRSPLQPLISLCEAPPSPLQLPGGNVTITYSYAGARAPMG  
QGFLLSYSQDWLMCLQEEFQCLNHRCSVSAVQRCDGVDACGDGSDEAGCSSDPFPGLTPRPVPS  
LPCNVTLEDFYGVFSSPGYTHLASVSHPOQSCHWLLDPHDGRRRLAVRFTALDLGFGDAVHVYDG  
PGPPESSRLRLSLTHFSNGKAVTVETLSGQAVVSYHTVAWSNGRGFNATYHVRGYCLPWDRPC  
GLGSGLGAGEGLGERCYSEAQRCDGSWDCADGTDEEDCPGCPGHPFGAAGTSGATACYLPA  
DRCNYQTFCADGADERRCRHCQPGNFRCRDEKCVYETWVCDGQPD CADGSDEWDCSYVLPRKV  
ITAAVIGSLVCGLLLVIALGCTCKLYAIRTQEYSIFAPLSRMEAEIVQQQAPPSYGQLIAQGA  
IPPVEDFPTENPNDNSVLGNLRSLLQILRQDMTPGGGPGARRRQGRRLMRRLVRRRLRWGLLP  
RTNTPARASEARSQVTPSAAPLEALDGGTGPAREGGAVGGQDGEQAPPLPIKAPLPSASTSPA  
PTTVPEAPGPLPSLPLEPSLLSGVVQALRGRLPSLGP PGPTRSPPGPHTAVLALEDEDDVLL  
VPLAEPGVWVAEAEDEPLLT

**Important features:****Signal peptide:**

amino acids 1-16

**Transmembrane domain:**

amino acids 442-462

**LDL-receptor class A (LDLRA) domain proteins**

amino acids 411-431, 152-171, 331-350 and 374-393

421/615

**FIGURE 417**

GTCTGTTCCCTTTGCTCTCTCGCGCCAGTCCTCCTCCCTGGTTCTCCTCAGCCGCTGTCGGAGGAGAGCACCCGGGA  
GACGCGGGCTGCAGTCGCGGGCGGCTTCTCCCGCCTGGGCGGCCTCGCCGCTGGGCAGGTGCTGAGCGCCCTAG  
AGCCTCCCTTGCCGCCTCCCTCCTCTGCCCCGGCCGACAGTCACATGGGGTGTGGAGGTAGATGGGCTCCCG  
GCCCCGGAGGGCGGCGGTGGATGCGGCGCTGGGCAGAACGAGCCGCGATTCCAGCTGCCCGCGCGCCCCGGGCG  
CCCCCTGCGAGTCCCCGCTTCAAGCCATGGGGACCTCTCCGAGCAGCAGCAGCCGCTCGCCTCCTGCAGCCGCATC  
GCCCCCGAGCCACAGCCACGATGATCGCGGGCTCCCTTCTCCTGCTTGGATTCCCTTAGCACCACCACAGCTCAG  
CCAGAACAGAGGGCCTCGAATCTCATTGGCACATACCGCCATGTTGACCGTGCCACCGGCCAGGTGCTAACCTGT  
GACAAGTGTCCAGCAGGAACCTATGTCTCTGAGCATTGTACCAACACAAGCCTGCGCGTCTGCAGCAGTTGCCCT  
GTGGGGACCTTTACCAGGCATGAGAATGGCATAGAGAAATGCCATGACTGTAGTCAGCCATGCCCATGGCCAATG  
ATTGAGAAATTACCTTGTGCTGCCTTGACTGACCGAGAATGCACTTGCCACCTGGCATGTTCCAGTCTAACGCT  
ACCTGTGCCCCCATACGGTGTGTCTGTGGGTGGGGTGTGCGGAAGAAAGGGACAGAGACTGAGGATGTGCGG  
TGTAAGCAGTGTGCTCGGGGTACCTTCTCAGATGTGCTTCTAGTGTGATGAAATGCAAAGCATACACAGACTGT  
CTGAGTCAGAACCTGGTGGTGATCAAGCCGGGGACCAAGGAGACAGACAACGTCTGTGGCACACTCCCGTCTTTC  
TCCAGCTCCACCTCACCTTCCCCTGGCACAGCCATCTTTCACGCCCCGAGCACATGGAAACCCATGAAGTCCCT  
TCCTCCACTTATGTTCCCAAAGGCATGAACCTCAACAGAATCCAACTCTTCTGCCTCTGTTAGACCAAAGGTACTG  
AGTAGCATCCAGGAAGGGACAGTCCCTGACAACACAAGCTCAGCAAGGGGGAAGGAAGACGTGAACAAGACCCCTC  
CCAAACCTTCAGGTAGTCAACCACCAGCAAGGCCCCACACAGACACATCCTGAAGTGTGCGGTCCATGGAG  
GCCACTGGGGGCGAGAAGTCCAGCAGCGCCATCAAGGGCCCCAAGAGGGGACATCCTAGACAGAACCTACACAAG  
CATTTTGACATCAATGAGCATTTGCCCTGGATGATTGTGCTTTTCTGCTGCTGGTGTGTTGGTGATTGTGGTG  
TGCAGTATCCGGAAGGCTCGAGGACTCTGAAAAAGGGGCCCCGGCAGGATCCAGTGCCATTGTGGAAGGCA  
GGGCTGAAGAAATCCATGACTCCAACCCAGAACCGGGAGAAATGGATCTACTACTGCAATGGCCATGGTATCGAT  
ATCCTGAAGCTTGTAGCAGCCCCAAGTGGGAAGCCAGTGGAAAGATATCTATCAGTTTCTTTGCAATGCCAGTGAG  
AGGGAGGTTGCTGCTTTCTCCAATGGGTACACAGCCGACCAGAGCGGGCCTACGCAGCTCTGCAGCATGGACC  
ATCCGGGGCCCCGAGGCCAGCCTCGCCAGCTAATTAGCGCCCTGCGCCAGCACCAGGAGAAACGATGTTGTGGAG  
AAGATTCTGGGCTGATGGAAGACACACCCAGCTGGAAACTGACAAACTAGCTCTCCCGATGAGCCCCAGCCCCG  
CTTAGCCCCAGCCCCATCCCCAGCCCCAACGCGAAACTTGAGAATTCCGCTCTCCTGACGGTGGAGCCTTCCCCA  
CAGGACAAGAACAAGGGCTTCTTCGTGGATGAGTCGGAGCCCTTCTCCGCTGTGACTCTACATCCAGCGGCTCC  
TCCGCGCTGAGCAGGAACGGTTCCTTTATTACCAAAGAAAAGAGGACACAGTGTGCGGCAGGTACGCTGGAC  
CCCTGTGACTTGCAGCCTATCTTTGATGACATGCTCCACTTTCTAAATCCTGAGGAGCTGCGGGTGATTGAAGAG  
ATCCCCAGGCTGAGGACAACTAGACCGGCTATTGGAATTTATGGAGTCAAGAGCCAGGAAGCCAGCCAGACC  
CTCCTGGACTCTGTTTATAGCCATCTTCTGACCTGCTGTAGAACATAGGGATACTGCATTCTGGAAATTACTCA  
ATTTAGTGGCAGGGTGGTTTTTTAATTTTCTTCTGTTTCTGATTTTTGTTGTTTGGGGTGTGTGTGTGTGTGTGT  
GT  
TCTCTCTCTTTTTTTTTTTTAAATAACTCTTCTGGGAAGTTGGTTTATAAGCCTTTGCCAGGTGTAACCTGTGTGAA  
ATACCCACCACTAAAGTTTTTTAAGTTCATATTTTCTCCATTTTGCCTTCTTATGTATTTTCAAGATTATCTG  
TGCACTTTAAATTTACTTAACTTACCATAAATGCAGTGTGACTTTTCCACACACTGGATTGTGAGGCTCTTAAC  
TTCTTAAAGTATAATGGCATCTTGTGAATCCTATAAGCAGTCTTTATGTCTCTTAACATTACACCTACTTTTT  
AAAAACAAATATTATTACTATTTTTATTATTGTTTGTCTTTATAAATTTTCTTAAAGATTAAGAAAATTTAAGA  
CCCCATTGAGTTACTGTAATGCAATCAACTTTGAGTTATCTTTTAAATATGTCTTGTATAGTTCAATATTCATGG  
CTGAACTTGACCACACTATTGCTGATTGTATGGTTTTTACCTGGACACCGTGTAGAATGCTTGATTACTTGTAC  
TCTTCTTATGCTAATATGCTCTGGGCTGGAGAAATGAAATCCTCAAGCCATCAGGATTTGCTATTTAAGTGGCTT  
GACAACCTGGGCCACCAAAGAACTTGAACCTTACCTTTTAGGATTTGAGCTGTTCTGGAACACATTGCTGCACCTT  
GGAAAGTCAAATCAAGTGCCAGTGGCGCCCTTTCCATAGAGAATTTGCCAGCTTTGCTTTAAAGATGTCTTG  
TTTTTTATATACATAATCAATAGGTCCAATCTGCTCTCAAGGCCTTGGTCTGTTGGGATTCTTCCACCAATT  
ACTTTAATTAATAATGGCTGCAACTGTAAAGACCCCTTGTCTGATATATTGCAACTATGCTCCCATTTACAAATG  
TACCTTCTAATGCTCAGTTGCCAGGTTCCAATGCAAAGGTGGCGTGGACTCCCTTTGTGTGGGTGGGGTTTGTGG  
GTAGTGGTGAAGGACCGATATCAGAAAAATGCCTTCAAGTGTACTAATTTATTAATAAACATTAGGTGTTTGTGA  
AAAAA



422/615

**FIGURE 418**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA52594

&gt;&lt;subunit 1 of 1, 655 aa, 1 stop

&gt;&lt;MW: 71845, pI: 8.22, NX(S/T): 8

MGTSPSSSTALASCSRIARRATATMIAGSLLLLLGFLSTTTAQPEQKASNLIQTYRHVDRATGQ  
VLTCDKCPAGTYVSEHCTNTSLRVCSSCPVGTFTRHENGIEKCHDCSQPCPWPMEIKLPCAAL  
TDRECTCPPGMFQSNATCAPHTVCPVGWGVRRKKGTTETEDVRCKQCARGTFSDVPSSVMKCKAY  
TDCLSQNLVVIKPGTKETDNVCGTLPFSFSSSTSPSPGTAFPRPEHMETHEVPSSTYVPKGMN  
STESNSSASVRPKVLSSIQEGTVPDNTSSARGKEDVNKTLPNLQVVNHQQGPHHRHILKLLPS  
MEATGGEKSSTPIKGPKRGHPRQNLHKHFDINEHLPWMIVLFLLLVLVIVVCSIRKSSRTLK  
KGPRQDPSAIVEKAGLKKSMTPTQNREKWIYYCNGHGIDILKLVAQVGSQWKDIYQFLCNAS  
EREVAAFSNGYTADHERAYAAALQHWRTIRGPEASLAQLISALRQHRRNDVVEKIRGLMEDTTQL  
ETDKLALPMSPSPSPSPSPNAKLENSALLTVEPSPQDKNKGFFVDESEPLLRC DSTSSGS  
SALSRNGSFITKEKKDTVLRQVRLDPCDLQPIFDDMLHFLNPEELRVIEEIPQAEDKLDRLFE  
IIGVKSQEASQTLLDSVYSHLPDLL

**Signal sequence:**

amino acids 1-41

**Transmembrane domain:**

amino acids 350-370

423/615

**FIGURE 419**

ATGGCTGGTGACGGCGGGGCGGGCAGGGGACCGGGGCGGGCCCGGGAGCGGGCCAGCTGCCGGGAGCCCTGA  
ATCACC GCCTGGCCCCGACTCCACCATGAACGTCGCGCTGCAGGAGCTGGGAGCTGGCAGCAACGTGGGATTCCAG  
AAGGGGACAAGACAGCTGTTAGGCTCACGCACGCAGCTGGAGCTGGTCTTAGCAGGTGCCTCTCTACTGCTGGCT  
GCACTGCTTCTGGGCTGCCTTGTGGCCCTAGGGGTCCAGTACCACAGAGACCCATCCCACAGCACCTGCCTTACA  
GAGGCCTGCATTCGAGTGGCTSGAAAAATCCTGGAGTCCCTGGACCGAGGGGTGAGCCCCCTGTGAGGACTTTTAC  
CAGTTCTCCTGTGGGGGCTGGATTTCGGAGGAACCCCTGCCGATGGGCGTTCTCGCTGGAAACACCTTCAACAGC  
CTCTGGGACCAAAACCAGGCCATACTGAAGCACCTGCTTGAAAACACCACCTTCAACTCCAGCAGTGAAGCTGAG  
CAGAAGACACAGCGCTTCTACCTATCTTGCCTACAGGTGGAGCGCATTGAGGAGCTGGGAGCCAGCCACTGAGA  
GACCTCATTGAGAAGATTGGTGGTTGGAACATTACGGGGCCCTGGGACCAGGACAACCTTTATGGAGGTGTTGAAG  
GCAGTAGCAGGGACCTACAGGGCCACCCATTCTTCACCGTCTACATCAGTGCCGACTCTAAGAGTTCCAACAGC  
AATGTTATCCAGGTGGACCACTGCTGGGCTCTTTCTGCCCTCTCGGGATTACTACTTAAACAGAACTGCCAATGAG  
AAAGTGCTCACTGCCTATCTGGATTACATGGAGGAACCTGGGGATGCTGCTGGGTGGGCGGCCACCTCCACGAGG  
GAGCAGATGCAGCAGGTGCTGGAGTTGGAGATACAGCTGGCCAACATCACAGTGCCCCAGGACCAGCGCGCGAC  
GAGGAGAAGATCTACCACAAGATGAGCATTTCGGAGCTGCAGGCTCTGGCGCCCTCCATGGACTGGCTTGAGTTC  
CTGTCTTTCTTGTGTCAACATTGGAGTTGAGTGACTCTGAGCCTGTGGTGGTGTATGGGATGGATTATTTGCAG  
CAGGTGTGAGAGCTCATCAACCGCACGGAACCAAGCATCCTGAACAATTACCTGATCTGGAACCTGGTGCAAAAG  
ACAACCTCAAGCCTGGACCGACGCTTTGAGTCTGCACAAGAGAAGCTGCTGGAGACCCCTCTATGGCACTAAGAAG  
TCCTGTGTGCCGAGGTGGCAGACCTGCATCTCCAACACGGATGACGCCCTTGCTTTGCTTTGGGGTCACTCTTC  
GTGAAGGCCACGTTTGACCGGCAAGCAAAGAAATTGCAGAGGGGATGATCAGCGAAATCCGGACCGCATTGTAG  
GAGGCCCTGGGACAGCTGGTTTGGATGGATGAGAAGACCCGCCAGGCAGCCAAGGAGAAAGCAGATGCCATCTAT  
GATATGATTGGTTTCCCAGACTTTATCCTGGAGCCCAAAGAGCTGGATGATGTTTATGACGGGTACGAAATTTCT  
GAAGATTCTTTCTTCCAAAACATGTTGAATTTGTACAACTTCTCTGCCAAGGTTATGGCTGACCAGCTCCGCAAG  
CCTCCCAGCCGAGACCAGTGGAGCATGACCCCCCAGACAGTGAATGCCTACTACCTTCCAACTAAGAAATGAGATC  
GTCTTCCCCGCTGGCATCCTGCAGGCCCCCTTCTATGCCCGCAACCAACCCCAAGGCCCTGAACCTTCGGTGGCATC  
GGTGTGGTCATGGGCCATGAGTTGACGCATGCCTTTGATGACCAAGGGCGAGTATGACAAAGAAGGGAACCTG  
CGGCCCTGGTGGCAGAATGAGTCCCTGGCAGCCTTCCGGAACACACGGCCTGCATGGAGGAACAGTACAATCAA  
TACCAGGTCAATGGGGAGAGGCTCAACGGCCGCCAGACGCTGGGGGAGAACATTACTGACAACGGGGGGCTGAAG  
GCTGCCTACAATGCTTACAAAGCATGGCTGAGAAAGCATGGGGAGGAGCAGCAACTGCCAGCCGTGGGGCTCACC  
AACCACCAGCTCTTCTTCGTGGGATTTGCCAGGTGTGGTGTCTCGGTCCGCACACCAGAGAGCTCTCACGAGGGG  
CTGGTGACCGACCCCCACAGCCCTGCCCGCTTCCGCGTGCTGGGCACCTCTCTCCAACCTCCCGTGACTTCTGCGG  
CACTTCGGCTGCCCTGTGCGCTCCCCCATGAACCCAGGGCAGCTGTGTGAGGTGTGGTTAGACCTGGATCAGGGGA  
GAAATGGCCAGCTGTACACAGACCTGGGGCAGCTCTCCTGACAAAGCTGTTTGCTCTTGGGTGGGAGGAAGCAA  
ATGCAAGCTGGGCTGGGTCTAGTCCCTCCCCCCCACAGGTGACATGAGTACAGACCCTCCTCAATCACCACATTG  
TGCTCTGCTTTGGGGGTGCCCTGCCTCCAGCAGAGCCCCCACCATTCACTGTGACATCTTCCGTGTACCCCT  
GCCTGGAAGAGGTCTGGGTGGGGAGGCCAGTTCCTATAGGAAGGAGTCTGCC

424/615

**FIGURE 420**

MNVALQELGAGSNVGFQKGTROQLLGSRTQLELVLAGASLLLAALLLGCLVALGVQYHRDPSHS  
TCLTEACIRVAGKILES LDRGVSPCEDFYQFSCGGWIRRNPLPDGRSRWNTFNSLWDQNQAIL  
KHLENTTFNSSEAEQKTQRFYLSCLQVERIEELGAQPLRDLIEKIGGWNITGPWDQDNFME  
VLKAVAGTYRATPFFTVYISADSKSSNSNVIQVDQSGFLPSRDYYLNRTANEKVLTA YLDYM  
EELGMLLGGRPTSTREQMQQVLELEIQ LANITVPQDQRRDEEKIYHKMSISELQALAPSMDWL  
EFLSFLLSPLLELSDSEPVVVYGMDYLQQVSELINRTEPSILNNYLIWNLVQKTTSSLD RRFES  
AQEKLETTYGTTKSCVPRWQTCISNTDDALGFALGSLFVKATFDRQSKEIAEGMISEIRTAF  
EEALGQLVWMDEKTRQAAKEKADAIYDMIGFPDFILEPKELDDVYDGYEISEDSFFQNM LNL  
NFSKVMADQLRKPPSRDQWSMTPQTVNAYYLP TKNEIVFPAGILQAPFYARNHPKALNFGGI  
GVVMGHELTHAFDDQGREYDKEGNLRPWWQNESLAAFRNHTACMEEQYNQYQVNGERLNGRQT  
LGENITDNGGLKAAYNAYKAWLRKHGEEQQLP AVGLTNHQLFFVGFAQVWCSVRTPESSHEGL  
VTDPHSPARFRVLGTLSNSRDFLRHFGCPVGSPMNP GQLCEVW

**Type II Transmembrane domain:**

amino acids 32-57

425/615

**FIGURE 421**

GGCGCCGCGTAGGCCCGGGAGGCCGGGCCGGGCTGCGAGCGCCTGCCCCATGCGCCGCC  
GCCTCTCCGCACG**ATG**TCCCCCTCGCGGAGGAAAGCGGCGCAGCTGCCCTGGGAGGACGGCAG  
GTCCGGGTTGCTCTCCGGCGGCCTCCCTCGGAAGTGTTCCGTCTTCCACCTGTTTCGTGGCCTG  
CCTCTCGCTGGGCTTCTTCTCCCTACTCTGGCTGCAGCTCAGCTGCTCTGGGGACGTGGCCCCG  
GGCAGTCAGGGGACAAGGGCAGGAGACCTCGGGCCCTCCCCGTGCCTGCCCCCAGAGCCGCC  
CCCTGAGCACTGGGAAGAAGACGCATCCTGGGGCCCCCACC GCCTGGCAGTGCTGGTGCCCTT  
CCGCGAACGCTTCGAGGAGCTCCTGGTCTTCGTGCCCCACATGCGCCGCTTCCTGAGCAGGAA  
GAAGATCCGGCACACATCTACGTGCTCAACCAGGTGGACCACTTCAGGTTCAACCGGGCAGC  
GCTCATCAACGTGGGCTTCCTGGAGAGCAGCAACAGCACGGACTACATTGCCATGCACGACGT  
TGACCTGCTCCCTCTCAACGAGGAGCTGGACTATGGCTTTCTGAGGCTGGGGCCCTTCACGT  
GGCCTCCCCGGAGCTCCACCCTCTCTACCACTACAAGACCTATGTCGGCGGCATCCTGCTGCT  
CTCCAAGCAGCACTACCGGCTGTGCAATGGGATGTCCAACCGCTTCTGGGGCTGGGGCCGCGA  
GGACGACGAGTTCTACCGGCGCATTAAAGGGAGCTGGGCTCCAGCTTTTCCGCCCCCTCGGGAAT  
CACAAC TGGGTACAAGACATTTCCGCCACCTGCATGACCCAGCCTGGCGGAAGAGGGACCAGAA  
GCGCATCGCAGCTCAAAAACAGGAGCAGTTCAAGGTGGACAGGGAGGGAGGCCTGAACACTGT  
GAAGTACCATGTGGCTTCCCGCACTGCCCTGTCTGTGGGCGGGGCCCCCTGCACTGTCTCAA  
CATCATGTTGGACTGTGACAAGACCGCCACACCCTGGTGCACATTCAGCT**TCAG**CTGGATGGAC  
AGTGAGGAAGCCTGTACCTACAGGCCATATTGCTCAGGCTCAGGACAAGGCCTCAGGTCGTGG  
GCCAGCTCTGACAGGATGTGGAGTGGCCAGGACCAAGACAGCAAGCTACGCAATTGCAGCCA  
CCCGGCCGCCAAGGCAGGCTTGGGCTGGGGCCAGGACACGTGGGGTGCCTGGGACGCTGCTTGC  
CATGCACAGTGATCAGAGAGAGGCTGGGGTGTGTCTGTCCGGGACCCCCCTGCCTTCCTGC  
TCACCCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGGTAGTGGGGAGGGCTGAACAGGA  
CAACCTCTCATCACCTACTCTGACCTCCTTCACGTGCCCAGGCCTGTGGGTAGTGGGGAGGG  
CTGAACAGGACAACCTCTCATCACCCCCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAA

426/615

**FIGURE 422**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56531
><subunit 1 of 1, 327 aa, 1 stop
><MW: 37406, pI: 9.30, NX(S/T): 1
MFPSRRKAAQLPWEDGRSGLLSGGLPRKCSVFHLFVACLSLGFFSLLWLQLSCSGDVARAVRG
QGQETSGPPRACPPPEPPPEHWEEDASWGPHRLAVLVPPFRERFEELLVFVPHMRRLSRKKIRH
HIYVLNQVDHFRFNRAALINVGFLSSNSTDYIAMHDVDLLPLNEELDYGFPFAGPFHVASPE
LHPLYHYKTYVGGILLLSKQHYRLCNGMSNRFWGWGREDDFYRRIKGAGLQLFRPSGITTGY
KTFRHLHDPAWRKRDKRIAAQKQEQFKVDREGGLNTVKYHVASRTALSVGGAPCTVLNIMLD
CDKTATPWCTFS
```

**Signal peptide:**

amino acids 1-42

**Transmembrane domain:**

amino acids 29-49 (type II)

**N-glycosylation site.**

amino acids 154-158

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 27-31

**Tyrosine kinase phosphorylation site.**

amino acids 226-233

**N-myristoylation site.**

amino acids 19-25, 65-71, 247-253, 285-291, 303-309, 304-310

427/615

**FIGURE 423**

CCATCCCTGAGATCTTTTTATAAAAAACCCAGTCTTTGCTGACCAGACAAAGCATAACCAGATC  
TCACCAGAGAGTCGCAGACACT**ATG**CTGCCTCCCATGGCCCTGCCAGTGTGTCCTGGATGCT  
GCTTTCTCCTGCCTCATTCTCCTGTGTGTCAGGTTCAAGGTGAAGAAACCCAGAAGGAAGTGCCTC  
TCCACGGATCAGCTGTCCCAAAGGCTCCAAGGCCTATGGCTCCCCCTGCTATGCCTTGTTTTT  
GTCACCAAAATCCTGGATGGATGCAGATCTGGCTTGCCAGAAGCGGCCCTCTGGAAAAGTGGT  
GTCTGTGCTCAGTGGGGCTGAGGGATCCTTCGTGTCCTCCCTGGTGAGGAGCATTAGTAACAG  
CTACTCATACATCTGGATTGGGCTCCATGACCCACACAGGGCTCTGAGCCTGATGGAGATGG  
ATGGGAGTGGAGTAGCACTGATGTGATGAATTACTTTGCATGGGAGAAAAATCCCTCCACCAT  
CTTAAACCCCTGGCCACTGTGGGAGCCTGTCAAGAAGCACAGGATTTCTGAAGTGGAAGATTA  
TAACTGTGATGCAAAGTTACCCTATGTCTGCAAGTTCAAGGACT**TAG**GGCAGGTGGGAAGTCAG  
CAGCCTCAGCTTGGCGTGCAGCTCATCATGGACATGAGACCAGTGTGAAGACTCACCTTGAA  
GAGAATATTCTCCCCAAACTGCCCTACCTGACTACCTTGTCATGATCCTCCTTCTTTTTCCTT  
TTTCTTCACCTTCATTTCAAGGCTTTTCTCTGTCTTCCATGTCTTGAGATCTCAGAGAATAATA  
ATAAAAATGTTACTTTATAAAAAAAAAAAAAAAAAAAAAA

428/615

**FIGURE 424**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA56965

&lt;subunit 1 of 1, 175 aa, 1 stop

&lt;MW: 19330, pI: 7.25, NX(S/T): 1

MLPPMALPSVSWMLLSCLILLCQVQGEETQKELPSPRISCPKGSKAYGSPCYALFLSPKSWMD

ADLACQKRPSGKLVSVLSGAEGSFVSSLVRSISNSYSYIWIGLHDPTQGSEPDGDGWEWSSTD

VMNYFAWEKNPSTILNPGHCGSLSRSTGFLKWKDYNCDAKLPYVCKFKD

**Important features:****Signal peptide:**

amino acids 1-26

**C-type lectin domain signature.**

amino acids 146-171

429/615

**FIGURE 425**

CGGACGCGTGGGCGCCACCTCCGGAACAAGCC**ATG**GTGGCGGCGACGGTGGCAGCGGCGTGG  
CTGCTCCTGTGGGCTGCGGCCTGCGCGCAGCAGGAGCAGGACTTCTACGACTTCAAGGCGGTC  
AACATCCGGGGCAAACCTGGTGTGCTGGAGAAGTACCGCGGATCGGTGTCCCTGGTGGTGAAT  
GTGGCCAGCGAGTGC GGCTTCACAGACCAGCACTACCGAGCCCTGCAGCAGCTGCAGCGAGAC  
CTGGGCCCCCACCACCTTTAACGTGCTCGCCTTCCCTGCAACCAGTTTGGCCAACAGGAGCCT  
GACAGCAACAAGGAGATTGAGAGCTTTGCCCCGCCGACCTACAGTGTCTCATTCCCCATGTTT  
AGCAAGATTGCAGTCACCGGTACTGGTGCCCATCCTGCCTTCAAGTACCTGGCCCAGACTTCT  
GGGAAGGAGCCCACCTGGAACCTTCTGGAAGTACCTAGTAGCCCCAGATGGAAAGGTGGTAGGG  
GCTTGGGACCCAACTGTGTCAGTGGAGGAGGTCAGACCCAGATCACAGCGCTCGTGAGGAAG  
CTCATCCTACTGAAGCGAGAAGACTT**TAA**CCACCGCGTCTCCTCCTCCACCACCTCATCCCG  
CCCACCTGTGTGGGGCTGACCAATGCAAACCTCAAATGGTGCTTCAAAGGGAGAGACCCACTGA  
CTCTCCTTCCTTTACTCTTATGCCATTGGTCCCATCATTTCTTGTGGGGGAAAAATTCTAGTAT  
TTTGATTATTTGAATCTTACAGCAACAAATAGGAACCTCCTGGCCAATGAGAGCTCTTGACCAG  
TGAATCACCAGCCGATACGAACGTCTTGCCAACAAAAATGTGTGGCAAATAGAAGTATATCAA  
GCAATAATCTCCCACCCAAGGCTTCTGTAAACTGGGACCAATGATTACCTCATAGGGCTGTTG  
TGAGGATTAGGATGAAATACCTGTGAAAGTGCCTAGGCAGTGCCAGCCAAATAGGAGGCATTC  
AATGAACATTTTTTGCATATAAACCAAAAAATAACTTGTTATCAATAAAAACTTGCATCCAAC  
ATGAATTTCCAGCCGATGATAATCCAGGCCAAAGGTTTAGTTGTTGTTATTTCTCTGTATTA  
TTTTCTTCATTACAAAAGAAATGCAAGTTCATTGTAACAATCCAAACAATACCTCACGATATA  
AAATAAAAATGAAAGTATCCTCCTCAAAAA



430/615

**FIGURE 426**

MVAATVAAAWLLLWAAACAQQEQDFYDFKAVNIRGKLVSLEKYRGSVSLVVNVASECGFTDQH  
YRALQQQLQORDLGPHHFENVLAFPCNQFGQQEPDSNKEIESFARRTYSVSFPMFSKIAVTGTGAH  
PAFKYLAQTSGKEPTWNFWKYL VAPDGKVVGAWDPTVSVEEVRPQITALVRKLILLKREDL

431/615

**FIGURE 427**

CAGTTCTGAAATCAATGGAGTTAATTTAGGGAATACAAACCAGCC**ATG**GGGGTGGAGATTGCC  
TTTGCCTCAGTGATTCTCACCTGCCTCTCCCTTCTGGCAGCAGGAGTCTCCCAGGTTGTTCTT  
CTCCAGCCAGTTCCAACCTCAGGAGACAGGTCCCAAGGCCATGGGAGATCTCTCCTGTGGCTTT  
GCCGGCCACTCA**TGA**GAGTGTTTTTGTGTAAAGTATTTTTTAGAATACTGTTGACTTCTTCAT  
GATTTAATAACCATCCTTTGCGAAGTTTTATGAGGCTTTAGGGGAATGTCAACCCTCAAATTT  
TTGTTATACTAGATGGCTTCCATTTACCCACCACTATTTTAAGGTCCCTTTATTTTTAGGTTT  
AAGGTTCAATTTGACTTGAGAAAGTGCCCTTCTGCAGCTTCATTGATTTTGTTTATCTTCACTA  
TTAATTGTAACGATTAAAAAAGAATAAGAGCACGCAGACCTCTAGGAGAATATTTTATCCCTG  
GGTGCCCTGACACATTTATGTAGTGATCCCACAAATGTGATTGTTAATTTAAATGTTATTCT  
AATATTAGTACATTCAGTTGTGATGTAATATGAATAACCAGAATCTATTTCTTAAAGTTTTG  
AGTATATTTTTCACTAGATATTTGTATAGAAAGACTGAATAGTGATG

432/615

**FIGURE 428**

MGVEIAFASVILTCLSLAAGVSQVLLQPVPTQETGPKAMGDLSCGFAGHS

433/615

**FIGURE 429**

CCAAAGTGATCATTTGAAAAAGAGATATCCACATCTTCAAGCCCATATAAAGGATAGAAGCTG  
CACAGGGCAGCTTTACTTACTCCAGCACCTTCCTCTCCCAGGCAA**ATG**GTGCTGACCATCTTT  
GGGATACAATCTCATGGATACGAGGTTTTTAAACATCATCAGCCCAAGCAACAATGGTGGCAAT  
G TTCAGGAGACAGTGACAATTGATAATGAAAAAATAACGCCATCGTTAACATCCATGCAGGA  
TCATGCTCTTCTACCACAATTTTTGACTATAAACATGGCTACATTGCATCCAGGGTGCTCTCC  
CGAAGAGCCTGCTTTATCCTGAAGATGGACCATCAGAACATCCCTCCTCTGAACAATCTCCAA  
TGGTACATCTATGAGAAACAGGCTCTGGACAACATGTTCTCCAACAATAACACCTGGGTCAAG  
TACAACCTCTGGAGTCTCTGATCAAAGACGTGGATTGGTTCCTGCTTGGGTCACCCATTGAG  
AAACTCTGCAAACATATCCCTTTGTATAAGGGGGAAGTGGTTGAAAACACACATAATGTCGGT  
GCTGGAGGCTGTGCAAAGGCTGGGCTCCTGGGCATCTTGGGAATTTCAATCTGTGCAGACATT  
CATGTT**TAG**GATGATTAGCCCTCTTGTTTTATCTTTCAAAGAAATACATCCTTGGTTTACAC  
TCAAAAGTCAAATTAAATTCTTCCCAATGCCCCAACTAATTTTGAGATTGAGTCAGAAAATA  
TAAATGCTGTATTTATA

434/615

**FIGURE 430**

><ss.DNA57834  
><subunit 1 of 1, 176 aa, 1 stop  
><MW: 19616, pI: 7.11, NX(S/T): 0  
MVLTIIFGIQSHGYEVFNIIISPSNNGGNVQETVTIDNEKNTAIVNIHAGSCSSTTIFDYKH  
GYIASRVLSRRACFILKMDHQNIPPLNNLQWYIYEKQALDNMFSNKYTWVKYNPLESLIK  
DVDWFLLGSPIEKLCKHIPLYKGEVVENTHNVGAGGCAKAGLLGILGISICADIHV

**Important features:****Signal peptide:**

Amino acids 1-26

**N-myristoylation sites:**

Amino acids 48-54;153-159;156-162;167-173

435/615

**FIGURE 431**

GCGTGGGG**ATG**TCTAGGAGCTCGAAGGTGGTGGCTGGGCCTCTCGGTGCTGCTGACGGCGGCCA  
CAGTGGCCGGCGTACATGTGAAGCAGCAGTGGGACCAGCAGAGGCTTCGTGACGGAGTTATCA  
GAGACATTGAGAGGCAAATTCGGAAAAAAGAAAACATTCGTCTTTTGGGAGAACAGATTATTT  
TGACTGAGCAACTTGAAGCAGAAAAGAGAGAAGATGTTATTGGCAAAAGGATCTCAAAAATCA**T**  
**GA**CTTGAATGTGAAATATCTGTTGGACAGACAACACGAGTTTGTGTGTGTGTGTTGATGGAGA  
GTAGCTTAGTAGTATCTTCATCTTTTTTTTTTGGTCACTGTCCTTTTAAACTTGATCAAATAAA  
GGACAGTGGGTCATATAAGTTACTGCTTTCAGGGTCCCTTATATCTGAATAAAGGAGTGTGGG  
CAGACACTTTTTTGGAAGAGTCTGTCTGGGTGATCCTGGTAGAAGCCCCATTAGGGTCACTGTC  
CAGTGCTTAGGGTTGTTACTGAGAAGCACTGCCGAGCTTGTGAGAAGGAAGGGATGGATAGTA  
GCATCCACCTGAGTAGTCTGATCAGTCGGCATGATGACGAAGCCACGAGAACATCGACCTCAG  
AAGGACTGGAGGAAGGTGAAGTGGAGGGAGAGACGCTCCTGATCGTCGAATCC

436/615

**FIGURE 432**

MSRSSKVVVLGLSVLLTAATVAGVHV KQQWDQQRLRDGVIRDIERQIRKKENIRLLGEQIILTE  
QLEAEREKMLLAKGSQKS

437/615

**FIGURE 433**

GAATTCGTGTCTCGGCACTCACTCCCGGCCGCCGGACAGGGAGCTTTCGCTGGCGCGCTTGGCCGGCGACAGGA  
CAGGTTCCGGGACGTCCATCTGTCCATCCGTCCGGAGAGAAATTACAGATCCGCAGCCCCGGGATGGGGCCGGCCCC  
CGCTGCCGCTGCTGCTGGGCCTCTTCTCCCGCGCTCTGGCGTAGAGCTATCACTGAGGCAAGGGAAGAAGCCA  
AGCCTTACCCGCTATTCCCGGGACCTTTTCCAGGGAGCCTGCAAACCTGACCACACACCGCTGTTATCCCTTCTC  
ACGCCAGTGGGTACCAGCCTGCCTTGATGTTTTACCAACCCAGCCTGGAAGACCACATACAGGAAACGTAGCCATT  
CCCCAGGTGACCTCTGTGAATCAAAGCCCCCTACCGCCTCTTGCCCTTCAAACACACAGTTGGACACATAATACTT  
TCTGAACATAAAGGTGTCAAATTTAATTGCTCAATCAATGTACCTAATATATACCAGGACACCACAATTTCTTGG  
TGGAAAGATGGGAAGGAATTGCTTGGGGGACATCATCGAATTACACAGTTTTATCCAGATGATGAAGTTACAGCA  
ATAATCGCTTCCCTCAGCATAACCAGTGTGCAGCGTTTACAGACAATGGTTCGTATATCTGTAAGATGAAATAAAC  
AATGAAGAGATCGTGTCTGATCCCATCTACATCGAAGTACAAGGACTTCTCACTTTACTAAGCAGCCTGAGAGC  
ATGAATGTCACCAGAAACACAGCCTTCAACCTCACCTGTCAGGCTGTGGGCCCGCCTGAGCCCGTCAACATTTTC  
TGGGTTCAAACAGTAGCCGTGTTAACGAACAGCCTGAAAAATCCCCCGGCGTGCTAACTGTTCCAGGCCCTGACG  
GAGATGGCGGTCTTCAGTTGTGAGGCCACAAATGACAAAGGGCTGACCGTGTCCAGGGAGTGACATCAACATC  
AAAGCAATTCCTCCCCACCAACTGAAGTCAGCATCCGTAACAGCACTGCACACAGCATTCTGATCTCCTGGGTT  
CCTGGTTTTGATGGATACTCCCCGTTCAAGGAATTGCAGCATTAGGTCAAGGAAGCTGATCCGCTGGGTAATGGC  
TCAGTCATGATTTTAAACACCTCTGCCTTACCACATCTGTACCAAATCAAGCAGCTGCAAGCCCTGGCTAATTAC  
AGCATTGGTGTTTTCTGCATGAATGAAATAGGCTGGTCTGCAGTGAGCCCTTGGATTCTAGCAAGCAGACTGAA  
GGAGCCCCATCAGTAGCACCTTAAATGTCACTGTGTTTTCTGAATGAATCTAGTGATAATGTGGACATCAGATGG  
ATGAAGCCTCCGACTAAGCAGCAGGATGGAGAAGTGGTGGGCTACCGGATATCCACGTGTGGCAGAGTGACAGGG  
ATTTCCAAAGAGCTCTTGGAGGAAGTTGGCCAGAATGGCAGCCGAGCTCGGATCTCTGTTCAAGTCCACAATGCT  
ACGTGCACAGTGAGGATTGCAGCCGTCAACAGAGGGGAGTTGGGCCCTTCACTGATCCAGTGAATAATTTATC  
CCTGCACACGGTTGGGTAGATTATGCCCCCTCTTCAACTCCGGCGCCTGGCAACGCAGATCCTGTGCTCATCATC  
TTTGGCTGCTTTTTGTGGATTATTTTGTATTGGGTTGATTTTATACATCTCCTTGGCCATCAGAAAAAGAGTCCAG  
GAGACAAAGTTTGGGAATGCATTACAGAGGAGGATTCTGAATTAGTGGTGAATTATATAGCAAAGAAATCCTTC  
TGTCGGCGAGCCATTGAACCTTACCTTACATAGCTTGGGAGTCAGTGAGGAACTACAAAATAAAGTGAAGATGTT  
GTGATTGACAGGAATCTTCTAATTTCTTGGAAAAATTTCTGGGTGAAGGAGAGTTTGGGTCTGTAATGGAAGGAAAT  
CTTAAGCAGGAAGATGGGACCTCTCTGAAAGTGGCAGTGAAGACCATGAAGTTGGACAACCTTTCACATCGGGAG  
ATCGAGGAGTTTCTCAGTGAGGCAGCGTGCATGAAGACTTTCAGCCACCCAAATGTCATTGCACTTCTAGGTGTG  
TGATAGAAATGAGCTCTCAAGGCATCCCAAAGCCCATGGTAATTTTACCCTTCATGAAATACGGGGACCTGCAT  
ACTTACTTACTTTATTTCCCGATTGGAGACAGGACCAAAGCATATTTCTCTGCAGACACTATTGAAGTTCATGGTG  
GATATTGCCCTGGGAATGGAGTATCTGAGCAACAGGAATTTTCTTCATCGAGATTTAGCTGCTCGAAACTGCATG  
TTGCGAGATGACATGACTGTCTGTGTTGCGGACTTCGGCCTCTCTAAGAAGATTTACAGTGGCGATTATTACCGC  
CAAGGCCGCATTGCTAAGATGCCTGTTAAATGGATCGCCATAGAAAGTCTTGCAGACCGAGTCTACACAAGTAAA  
AGTGATGTGTGGGCATTGGCGTGACCATGTGGGAAATACGTACGCGGGGAATGACTCCCTATCTTGGGGTCCAG  
AACCATGAGATGTATGACTATCTTCTCCATGGCCACAGGTTGAAGCAGCCCCGAAGACTGCCTGGATGAACTGTAT  
GAAATAATGTACTCTTGCTGGAGAACCGATCCCTTAGACCGCCCCACCTTTTCACTATTGAGGCTGCAGCTAGAA  
AACTCTTAGAAAGTTTGCCTGACGTTCCGAACCAAGCAGACGTTATTTACGTCAATACACAGTTGCTGGAGAGC  
TCTGAGGGCCTGGCCCAGGGCCCCACCTTGCTCCACTGGACTTGAACATCGACCCCTGACTCTATAATTGCCTCC  
TGCACTCCCCGCGCTGCCATCAGTGTGGTCACAGCAGAAGTTCATGACAGCAAACCTCATGAAGGACGGTACATC  
CTGAATGGGGGCGAGTGAGGAATGGGAAGATCTGACTTCTGCCCCCTCTGCTGCAGTCACAGCTGAAAAGAACAGT  
GTTTTACCGGGGGAGAGACTTGTTAGGAATGGGGTCTCCTGGTCCCATTGAGCATGCTGCCCTTGGGAAGCTCA  
TTGCCCGATGAACTTTTGTGTTGCTGACGACTCCTCAGAAGGCTCAGAAGTCTGATGTGAGGAGAGGTGCGGGGA  
GACATTCCAAAAATCAAGCCAATTCTTCTGCTGTAGGAGAATCCAATTGTACCTGATGTTTTTGGTATTTGTCTT  
CCTTACCAAGTGAACCTCCATGGCCCCAAAGCACCAGATGAATGTTGTTAAGGAAGCTGTCATTAATAAATACATAA  
TATATATTTATTTAAAGAGAAAAATATGTGTATATCATGAAAAAGACAAGGATATTTTAATAAAACATTACTTA  
TTTCATTTCACTTATCTTGCATATCTTAAATTAAGCTTCAGCTGCTCCTTGATATTAACCTTTGTACAGAGTTG  
AAGTTGTTTTTCAACTTCTTTTCTTTTCACTATTAATAATGTAATAAATATTTGTAAATGAAATGCCATATT  
TGACTTGGCTTCTGGTCTTGATGTATTTGATAAGAATGATTAATTTTCTGATATGGCTTCCATAATAAATTGAA  
ATAGGA



438/615

**FIGURE 434**

MGPAPLPLLLGLFLPALWRRRAITEAREEAKPYPLFPGPFGSLQTDHTPLLSLPHASGYQPALMFSPTQPGRPHT  
GNVAIPQVTSVESKPLPPLAFKHTVGHIIILSEHKGVKFNCSINVPNYQDTTISWWKDGKELLGGHHRITQFYYP  
DEVTAIIASFSITSVQRSDNGSYICKMKINNEEIVSDPIYIEVQGLPHFTKQPESMNVTNTAFNLTCQAVGPPE  
PVNIFWVQNSSRVNEQPEKSPGVLTVPGLTEMAVFSCEAHNDKGLTVSQGVQINIKAIKPSPTTEVSIRNSTAHSI  
LISWVPGFDGYSPFRNCISIQVKEADPLNGSVMIFNTSALPHLYQIKQLQALANYSIGVSCMNEIGWSAVSPWIL  
ASTTEGAPSVAPLNVTVFLNESSDNVDIRWMKPPTKQQDGELVGYRISHVWQSAGISKELLEEVQNGSRARISV  
QVHNATCTVRIAAVTRGGVGPFSDPVKIFIPAHGWVDYAPSSTPAPGNADPVLIIFGCFCGFIILGLILYISLAI  
RKRVQETKFGNAFTEEDSELVVNYIAKKSFCRRRAIELTLHSLGVSEELQNKLEDVVIDRNLLILGKILGEGEFGS  
VMEGNLKQEDGTSKLVAVKTMKLDNSSHREIEEFLSEAACMKDFSHPNVIRLLGVCIEMSSQGIPKPMVILPFMK  
YGDHLTYLLYSRLETGPKHIPLQTLKFMVDIALGMEYLSNRNFLHRDLAARNCMLRDDMTVCVADFGLSKKIYS  
GDYYRQGRIAKMPVKWIAIESLADRVYTSKSDVWAFGVTMWEIRTRGMTYPYGVQNHMYDYLLHGHRLKQPEDC  
LDELYEIMYSCWRTPDLPDRPTFSVLRQLQLEKLLLESIPDVRNQADVIVNTQLLESSEGLAQGPPLAPLDLNDP  
SIIASCTPRAAISVVTAEVHDSKPHEGRYILNGGSEEWEDLTSAPSAAVTAEKNSVLPGERLVRNGVSWSHSSML  
PLGSSLPDELLFADDSSEGSEVLM

Signal sequence:	Amino acids 1-18
Transmembrane domain:	Amino acids 501-520
N-glycosylation sites:	Amino acids 114-118;170-174;207-211; 215-219;234-238;294-298;316-320;329-333; 336-340;354-358;389-393;395-399;442-446; 454-458;625-629
Tyrosine kinase phosphorylation sites:	Amino acids 675-683;865-873;923-930
N-myristoylation sites:	Amino acids 41-47;110-116;171-177; 269-275;275-281;440-446;507-513;535-541; 966-972
Prokaryotic membrane lipoprotein lipid attachment site:	Amino acids 351-362
Tyrosine protein kinases specific active-site signature:	Amino acids 719-732

439/615

**FIGURE 435**

AATGTGAGAGGGGCTGATGGAAGCTGATAGGCAGGACTGGAGTGTTAGCACCACTACTGGATG  
TGACAGCAGGCAGAGGAGCACTTAGCAGCTTATTCAGTGTCCGATTCTGATTCCGGCAAGGAT  
CCAAGC**ATG**GAATGCTGCCGTCGGGCAACTCCTGGCACACTGCTCCTCTTTCTGGCTTTCCTG  
CTCCTGAGTTCCAGGACCGCACGCTCCGAGGAGGACCGGGACGGCCTATGGGATGCCTGGGGC  
CCATGGAGTGAATGCTCACGCACCTGCGGGGGAGGGGCCTCCTACTCTCTGAGGCGCTGCCTG  
AGCAGCAAGAGCTGTGAAGGAAGAAATATCCGATACAGAACATGCAGTAATGTGGACTGCCCA  
CCAGAAGCAGGTGATTTCCGAGCTCAGCAATGCTCAGCTCATAATGATGTCAAGCACCATGGC  
CAGTTTTATGAATGGCTTCCTGTGTCTAATGACCCTGACAACCCATGTTCACTCAAGTGCCAA  
GCCAAAGGAACAACCCTGGTTGTTGAACTAGCACCTAAGGTCTTAGATGGTACGCGTTGCTAT  
ACAGAATCTTTGGATATGTGCATCAGTGGTTTATGCCAAATTGTTGGCTGCGATCACCAGCTG  
GGAAGCACCGTCAAGGAAGATAACTGTGGGGTCTGCAACGGAGATGGGTCCACCTGCCGGCTG  
GTCCGAGGGCAGTATAAATCCCAGCTCTCCGCAACCAAATCGGATGATACTGTGGTTGCACTT  
CCCTATGGAAGTAGACATATTGCGCTTGTCTTAAAGGTCTGATCACTTATATCTGGAAACC  
AAAACCCCTCCAGGGGACTAAAGGTGAAAACAGTCTCAGCTCCACAGGAACTTTCCTTGTGGAC  
AATTCTAGTGTGGACTTCCAGAAATTTCCAGACAAAGAGATACTGAGAATGGCTGGACCACTC  
ACAGCAGATTTTCATTGTCAAGATTCGTAACCTCGGGCTCCGCTGACAGTACAGTCCAGTTCATC  
TTCTATCAACCCATCATCCACCGATGGAGGGAGACGGATTTCTTTCCTTGCTCAGCAACCTGT  
GGAGGAGGTTATCAGCTGACATCGGCTGAGTGCTACGATCTGAGGAGCAACCGTGTGGTTGCT  
GACCAATACTGTCACTATTACCCAGAGAACATCAAACCCAAACCCAAGCTTCAGGAGTGCAAC  
TTGGATCCTTGTCCAGCCAGTGACGGATACAAGCAGATCATGCCTTATGACCTCTACCATCCC  
CTTCCTCGGTGGGAGGCCACCCCATGGACCGCGTGCTCCTCCTCGTGTGGGGGGGGCATCCAG  
AGCCGGGCAGTTTCCTGTGTGGAGGAGGACATCCAGGGGCATGTCACTTCAGTGGAAAGAGTGG  
AAATGCATGTACACCCCTAAGATGCCCATCGCGCAGCCCTGCAACATTTTTGACTGCCCTAAA  
TGGCTGGCACAGGAGTGGTCTCCGTGCACAGTGACATGTGGCCAGGGCCTCAGATACCGTGTG  
GTCCTCTGCATCGACCATCGAGGAATGCACACAGGAGGCTGTAGCCCAAAAACAAAGCCCCAC  
ATAAAAGAGGAATGCATCGTACCCACTCCCTGCTATAAAACCAAAGAGAACTTCCAGTCGAG  
GCCAAGTTGCCATGGTTCAAACAAGCTCAAGAGCTAGAAGAAGGAGCTGCTGTGTGTCAGAGGAG  
CCCTCG**TAA**GTTGTAAAAGCACAGACTGTTCTATATTTGAAACTGTTTTGTTTAAAGAAAGCA  
GTGTCTCACTGGTTGTAGCTTTCATGGGTTCTGAACTAAGTGTAATCATCTCACCAAAGCTTT  
TTGGCTCTCAAATTAAAGATTGATTAGTTTCAAAAAAAAAA

440/615

**FIGURE 436**

&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA58847

&lt;subunit 1 of 1, 525 aa, 1 stop

&lt;MW: 58416, pI: 6.62, NX(S/T): 1

MECCRRATPGTLLLFIAFLLLSSRTARSEEDRDGLWDAWGPWSECSRTC GGGASYSLRRCLSS  
KSCEGRNIRYRTCSNVDCPPEAGDFRAQQCSAHNDVKHHGQFYEWLPVSNDPDNPCSLKCQAK  
GTTLVVELAPKVLDTGTRCYTESLDMCISGLCQIVGCDHQLGSTVKEDNCGVCNGDGSTCRLVR  
GQYKSQLSATKSDDTVVALPYGSRHIRLVKGPDHLYLETKTLQGTKGENSLSSSTGTFLVDNS  
SVDFQKFDPKEILRMAGPLTADFIVKIRNSGSADSTVQFIFYQPIIHRWRETDFFP SATCGG  
GYQLTSAECYDLRSNRVVADQYCHYYPENIKPKPKLQECNLDPCPASDGYKQIMPYDLYHPLP  
RWEATPWTACSSSCGGGIQSRVSCVEEDIQGHVTSVEEWKCMYTPKMPIAQPCNIFDCPKWL  
AQEWSPTVTCGQGLRYRVLCIDHRGMHTGGCSPKTKPHIKEECIVPTPCYKPKEKLPVEAK  
LPWFKQAQELEEGAAVSEEPS

**Important features:****Signal peptide:**

amino acids 1-25

**N-glycosylation site.**

amino acids 251-254

**Thrombospondin 1**

amino acids 385-399

**von Willebrand factor type C domain proteins**

amino acids 385-399, 445-459 and 42-56

441/615

**FIGURE 437**

AACTGGAAGGAAAGAAAGAAAGGTCAGCTTTGGCCCAGATGTGGTTACCCCTTGGTCTCCTGT  
CTTTATGTCTTTCTCCTCTTCCTATTCTGTTCATCTCCCTCACTTAAGTCTCAGGCCTGTCAGC  
AGCTCCTGTGGACATTGCCATCCCCCTCTGGTAGCCTTCAGAGCAAACAGGACAACCTATGTTA  
TGGATGTTTCCACCAACCAGGGTAGTGGCATGGAGCACCGTAACCATCTGTGCTTCTGTGATC  
TCTATGACAGAGCCACTTCTCCACCTCTGAAATGTTCCCTGCTCTTGAAATCTGGCATGAGATG  
GCACAGGTGACCACGCAGAAGCCACCAGAATCTTGCTGCCCCTATTCCTCCTCCCAAGTCTGT  
TCTCTTATTGTCAACCTCAGCACAAACAGGCTGGCGCCAATGGCATTACAGAGAAAGCAATCTG  
TGTGGCTAGTGGGCAGATTACCATGCAAGCCCCAGGAGAAATGGAGGAGCTTTGTAGCCACCT  
CCCTGTCAGCCAGTATTAACATGTCCCCTTCCCCCTGCCCCGCCGTAGATTACAGGACATTCGC  
CCCTGTGTGCCACCAAACCAGGACTTTCCCCTTGGCTTGGCATCCCTGGCTCTCTCCTGGTAC  
CCAGCAAGACGTCTGTTCCAGGGCAGTGTAGCATCTTTCAAGCTCCGTTACTATGGCGATGGC  
CATGATGTTACAATCCCACCTTGCTGAATAATCAAGTGGGAAGGGGAAGCAGAGGGAAATGGG  
GCCATGTGAATGCAGCTGCTCTGTTCTCCCTACCCTGAGGAAAAACCAAAGGGAAGCAACAGG  
AACTTCTGCAACTGGTTTTTATCGGAAAGATCATCCTGCCTGCAGATGCTGTTGAAGGGGCAC  
AAGAAATGTAGCTGGAGAAGATTGATGAAAGTGCAGGTGTGTAAGGAAATAGAACAGTCTGCT  
GGGAGTCAGACCTGGAATTCTGATTCCAAACTCTTTATTACTTTGGGAAGTCACTCAGCCTCC  
CCGTAGCCATCTCCAGGGTGACGGAACCCAGTGTATTACCTGCTGGAACCAAGGAAACTAACA  
ATGTAGGTTACTAGTGAATACCCAATGGTTTCTCCAATTATGCCCATGCCACCAAAACAATA  
AAACAAAATTCTCTAACACTGAAA

442/615

**FIGURE 438**

MWLPLGLLSICLSPLPILSSPSLKSQACQQLWTLPSPLVAFRANRTTYVMDVSTNQSGMEH  
RNHLCFCDLYDRATSPPLKCSLL

**FIGURE 439**

TTTCTCATAGTTGGCGTCTTCTAAAGGAAAAAACTAAAATGAGGAACCTCAGCGGACCGGGAGCGACGCAGCTT  
GAGGGAAGCATCCCTAGCTGTTGGCGCAGAGGGGCGAGGCTGAAGCCGAGTGGCCCCAGGTTGCTGATGGGGCTGG  
GCGAAAGGTGAAAGAGTTTCAGAAACAAGCTTCTTGGAAACCCATGAGCTCTTGCGACATTATACCGT  
CTGAGGGTAGCAGCTCGAAACTAGAAGAAGTGGAGTGTTCAGGGACGGCAGTATCTCTTTGTGTGACCTGGC  
GGCCTATAGGGACGTTGGCTTCAGACCTTTGTGATACACCATGCTGCGTGGGACGATGACGGCGTGGAGAGGAATG  
AGGCCTGAGGTCACACTGGCTTGCTCTCTAGCCACAGCAGGCTGCTTTGCTGACTTGAACGAGGTCCCTCAG  
GTCAACCGTCCAGCTGCGTCCACCGTCCAGAAGCCCGAGGCATGTGATCTTGGGCTGCGTGGTGGAACTCCCA  
AGGATGAATGTAACTTGGCGCTGAATGGAAGGAGCTGAATGGCTCGGATGATGCTCTGGGTGCTCTCATCACC  
CAGGGGACCCCTCGTCATCACTGCCCTTAACAACCACACTGTGGGACGGTACCAGTGTGTGGCCCGGATGCCGCC  
GGGCTGTGGCCAGCGTGCCAGCCACTGTGACATAGCCAATCTCCAGGACCTCAAGTTAGATGTGCAGCAGCT  
ATTGAAGTGGATGAGGGAACACAGCTGCTATTGCTGCCACTGCTGAGAGCCACCCCAAGCCACGGTCCGG  
TACAGCGTCAAACAAGAGTGGCTGAGGCTCCAGAGGTAACCTACCTGATGCCCTCAGGGAACTCCAGATT  
GTGAATGCCAGCCAGGAGGACGAGGGCATGTACAAGTGTGCAGCCTACAACCCAGTGACCCAGGAAGTGA  
TCCGGCTCCAGGACAGGCTACGTGTGCGCGCTCCACCGCTGAGGCTGCCCGCATCATCTACCCCCAGAGGCC  
CAAACCATCATCGTCACCAAGGCGAGAGTCTATTCTGGAGTGTGTGGCCAGTGGAACTCCCAACCCCGGTC  
ACCTGGGCCAAGGATGGGTCCAGTGTACCGGTACACAAGAGCGCTTCTGCTGAGCAACCTCTCATCGAC  
ACCACCGCAGGAGGAGTCTAGGCACCTACCGCTGCATGGCCGACAATGGGGTTGGGCAGCCCGGGGCGAGCGGTC  
ATCCTCTACAATGTCCAGGTGTTTGAACCCCTGAGGTACCATGGAGCTATCCAGCTGGTATCCCCCTGGGG  
CAGAGTGCCAAGCTTACCTGTGAGGTGCGTGGGAACCCCGCCCTCCGCTGCTGTGGCTGAGGAATGCTGTGCC  
TTCATCTCCAGCGACGCTCCGGCTCTCCGCGAGGGCTCGCGCTGCTCAGCATGGGGCTGAGGACGAGGG  
GTCATTACAGTGCATGGCCGAGAACGAGGTTGGGAGCGCCCATGCCGTAGTCCAGCTGCGGACCTCCAGGCCAAGC  
ATAACCCCAAGGCTATGGCAGGATGCTGAGCTGGCTAGTGGCACACCTCCTGTATCACCTCCAACTCGGCAAC  
CCTGAGCAGATGCTGAGGGGGCAACCGGCGCTCCCCAGACCCCACTGCTCAGTGGGGCTGCTTCCCCGAGTGT  
CCAGGAGAGAAGGGGCGAGGGGCTCCGCGAGGCTCCCATCATCTCAGCTGCCCCGCACCTCCAGACGAGC  
TCATATGAACTGGTGTGGCGGCTCGGCATGAGGGCAGTGGCCGGGCGCCAATCTCTACTATGTGGTGAACAC  
CGCAAGCAGGTCAAAATTCCTGTACGATTGACCATCTCTGGCATCCAGCCCAACGACCGCCCTGACCTC  
ACGAGCTTGACCCCGGGAGCTTGTATGAATGGAGATGGCAGCTTACAACCTGTGCGGAGAGGGCCAGACGCC  
ATGGTCACTTCCGAACTGGACGGCGGCCAAACCCGAGATCATGGCACAAGAGCAGCATCCAGAGAGAC  
GACCTTGAGCCAGTCCCCAGAGCAGCAGCCAGCCAGACCACGGCCGCTCTCCCCCCAGAAGCTCCCGACAGG  
CCCACCATCTCCACGGCTCCGAGACCTCAGTGTACGTGACCTGATTCCCCGTGGGAATGGTGGGTTCCCAATC  
CAGTCTCTCCGTGGAGTACAAGAGCTAAGAAGTGGGAGACTGGATTCTGGCCACGACGCCATCCCCCA  
TCGGGGTGTCCGTGGAGATCAGGGCTAGAGAAGGCACTCCTACAAGTTTCAGTCCGGGCTGTGAACATG  
CTGGGGGAGAGCGAGGCCAGCGCCCCCTCTCGGCCCTACGTGGTGTCCGGCTACAGCGGTGCGGTGTACGAGAGG  
CCCGTGGCAGGTCTTATATCACTTACAGGATGCGGTCAATGAGACCACCTCATGCTCAAGTGGATGTACATC  
CCAGCAAGTAAACAACAACACCCCAATCCATGGCTTTTATATCTATTATCGACCCACAGACAGTACAATGATG  
GACTACAGAAGGATATGGTGAAGGGGACAAGTACTGGCACTCCTACGACCACCTGCAGCCAGAGACCTCCTAC  
GACATTAAGATGCAGTGCTTCAATGAAGGAGGGGAGAGCGAGTTCAGCAACGTGATGATCTGTGAGACCAAAGCT  
CGGAAGTCTTCTGGCCAGCCTGGTGCAGTGCACCCCACTCTGGCCCCACACAGCGCCCCCTTCTGAAACC  
ATAGAGCGGCGCGTGGGCACTGGGGCCATGGTGGCTCGCTCCAGCGACCTGCCCTATCTGATTGTCCGGGTGCT  
CTGGGTCCTATCTTCTCATCATCGTCACTTCTATCCCCCTCTGCTGTGGAGGGCTGGTCTAAGCAAAAACAT  
ACAACAGACCTGGGTTTTCTCGAAGTGCCTTCCACCTCCTGCCGTATACTATGGTGCCATTGGGAGGACTC  
CCAGGCCACCAGGCCAGTGGACAGCCCTACTCAGTGGCATCAGTGGACGGGCTGTGCTAATGGGATCCCATG  
AATAGGGGCTGCCCCCTCGGCTGCAGTGGGCTACCCGGGCATGAAGCCCCAGCAGCATGCCAAGCGAGCTTAC  
CAGCAGGTGACACACGAGCCTGCTGAGGCAGACCATCTTGGCAATGGATGATGACCCCAAGTCACCATG  
ACGAGGGGTCCCAAGTCTAGCCCGGACGAGGGCTCTTCTTATACACACTGCCCGACGACTCCACTCACCAGCTG  
CTGCAGCCCATCAGCATGCTGCCAACGCCAGGAGCAGCTGCTGCTGTGGGCGAGTCAGGGGTGAGGAGAGCC  
CCGACAGTCTCTGCTGGAAGCAGGTGGGACCTCCATTCTCACTCAGGCCCCCATGCTGCTTGGGCTTGTG  
CCAGTTGAAGAGGTGGACAGTCTGACTCTGCCAAGTGAGTGGAGGAGACTGGTGTCCCCAGCACCCCGTAGGG  
GCCTACGTAGGACAGGAACCTGGAATGCAGCTCTCCCCGGGGCCACTGGTGGTGTGTCTTTTGAACACCACCT  
CTCACAATTAGGCAGAAGCTGATATCCAGAAAGACTATATATTGTTTTTTTTTAAAAAAGAGAAAAA  
AGAGACAGAGAAAATTGGTATTTATTTTCTATTATAGCCATATTTATATTTATGCACTTGTAAATAATGTA  
TATGTTTTATAATTCTGGAGAGACATAAGGAGTCTACCCGTTGAGTTGGAGAGGAAAAAAGAAGCTGCCA  
CCTAACAGGAGTCACCCGAGAAAGCACCGCACAGGCTGGCGCGGGACAGACTCCTAACCTGGGGCCTCTGCAGTG  
GCAGGCGAGGCTGCAGGAGGCCACAGATAAGCTGGCAAGAGGAAGGATCCAGGCACATGGTTCATCAGGACA  
TGAGGGAAACAGAGGGGACGCGTATCACAGCCTGGAGACACCCACACAGATGGCTGGATCCGGTGCTACGGGAA  
ACATTTTCTAAGATGCCATGAGAACACAGCAAGATGTGTACAGCACTATGACATTAATAAACCTCCAGAAAT  
CAATAATCCGTGGCAACATATCTCTGTAAAAACAACACTGTAACTCTAATAAATGTTTAGTCTTCCCTGTAAAA

444/615

**FIGURE 440**

MLRGTM TAWRGMRPEVTLACL LLLATAGCFADLNEVPQVTVQPASTVQKPGGT VILGCVVEPPR  
MNVTWRLNGKELNGSDDALGVLI THGTLVIT ALNNHTVGRYQCVARMPAGAVASVPATVTLAN  
LQDFKLDVQH VIEVDEGNTAVIACHLPESH PKAQVRYSVKQEWLEASRGNYLIMPSGNLQIVN  
ASQEDEGMYKCAAYNPVTQE VKTSGSSDRLRVRRSTAEAA RIIYPPEAQTIIVTKGQSLILEC  
VASGIPPPRV TWAKDGSSVTGYNKTRFLLSNLLIDTTSEEDSGTYRCMADNGVGQPGA AVILY  
NVQVFEPPEVTMELS QLVI PWGQSAKLTCEVRGNPPPSVLWLRNAVPLISSQRLRLSRRALRV  
LSMGPEDEGVYQCMAENEV GSAHAVVQLRTSRPSITPRLWQDAELATGT PPVSPSKLGNPEQM  
LRGQPALPRPPTSVGPASPKCPGEGKGQGA PAEAPIILSSPRTSKTDSYELVWRPRHEGSGRAP  
ILYYVVKHRKQVTNSSDDWTISGIPANQHRLTLTRLDPGSLYEVE MAAYNCAGEGQTAMVTFR  
TGRRPKPEIMASKEQQIQRDDPGAS PQSSSQPDHGR LSPPEAPDRPTISTASETSVYVTWIPR  
GNGGFP IQSFRVEYKKLKKVGDWILATSAIPPSRLSVEITGLEKGT SYKFRVRALNMLGESEP  
SAPSRPYV VSGYSGRVYERPVAGPYITFTDAVN ETTIMLKWMI PASNNNTPIHGFYIYYRPT  
DSDNDS DYKKDMVEGD KYWHSISHLQ PETS YDIKMQC FNEGGESEFSNVMICETKARKSSGQP  
GRLPPPTLAPPQPPLPETIERPVGTGAMVARSSDLPYLIVGVVLG SIVLIIVTFIPFCLWRAW  
SKQKHTTDLGFPRSALPPSCP YTMVPLGGLPGHQASGQPYLSGISGRACANGIHMNRGCPSAA  
VGYPGMKPQQHCPGELQQQSDTSSLLRQTHLGNGYDPQSHQITRGPKSSPDEGSFLYTL PDDS  
THQLLQPHHDCCQRQE QPAAVGQSGVRRAPDSPVLEAVWD PPFHSGPPCCLGLVPVEEVDSPD  
SCQVSGGDWCPQH PVGAYVGQEPGMQLSPGPLVRVSFETPPLTI

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 16-30 (type II), 854-879

445/615

**FIGURE 441**

GAGAGAATAGCTACAGATTCTCCATCCTCAGTCTTTGCAAGGCGACAGCTGTGCCAGCCGGGC  
TCTGGCAGGCTCCTGGCAGC**ATG**GCAGTGAAGCTTGGGACCCTCCTGCTGGCCCTTGCCCTGG  
GCCTGGCCCAGCCAGCCTCTGCCCCGCCGGAAGCTGCTGGTGTTTCTGCTGGATGGTTTTCGCT  
CAGACTACATCAGTGATGAGGCGCTGGAGTCATTGCCTGGTTTCAAAGAGATTGTGAGCAGGG  
GAGTAAAAGTGGATTACTTGACTCCAGACTTCCCTAGTCTCTCGTATCCCAATTATTATACCC  
TAATGACTGGCCGCCATTGTGAAGTCCATCAGATGATCGGGAACACATGTGGGACCCACCA  
CCAACAAGTCCTTTGACATTGGCGTCAACAAAGACAGCCTAATGCCTCTCTGGTGGAATGGAT  
CAGAACCTCTGTGGGTCACTCTGACCAAGGCCAAAAGGAAGGTCTACATGTACTACTGGCCAG  
GCTGTGAGGTTGAGATTCTGGGTGTGAGACCCACCTACTGCCTAGAATATAAAAATGTCCCAA  
CGGATATCAATTTTGCCAATGCAGTCAGCGATGCTCTTGACTCCTTCAAGAGTGGCCGGGCCG  
ACCTGGCAGCCATATACCATGAGCGCATTGACGTGGAAGGCCACCACTACGGGCCTGCATCTC  
CGCAGAGGAAAGATGCCCTCAAGGCTGTAGACACTGTCCTGAAGTACATGACCAAGTGGATCC  
AGGAGCGGGGCCTGCAGGACCGCCTGAACGTCATTATTTTCTCGGATCACGGAATGACCGACA  
TTTTCTGGATGGACAAAGTGATTGAGCTGAATAAGTACATCAGCCTGAATGACCTGCAGCAAG  
TGAAGGACCGCGGGCCTGTTGTGAGCCTTTGGCCGGCCCCTGGGAAACACTCTGAGATATATA  
ACAAACTGAGCACAGTGGAACACATGACTGTCTACGAGAAAGAAGCCATCCCAAGCAGGTTCT  
ATTACAAGAAAGGAAAGTTTGTCTCTCCTTTGACTTTAGTGGCTGATGAAGGCTGGTTCATAA  
CTGAGAATCGAGAGATGCTTCCGTTTTGGATGAACAGCACCGGCAGGCGGGAAGGTTGGCAGC  
GTGGATGGCACGGCTACGACAACGAGCTCATGGACATGCGGGGCATCTTCCTGGCCTTCGGAC  
CTGATTTCAAATCCAACCTTCAGAGCTGCTCCTATCAGGTCGGTGGACGTCTACAATGTCATGT  
GCAATGTGGTGGGCATCACCCCGCTGCCCCAACAACGGATCCTGGTCCAGGGTGATGTGCATGC  
TGAAGGGCCGCGCCGGCACTGCCCCGCCTGTCTGGCCCAGCCACTGTGCCCTGGCACTGATTC  
TTCTCTTCCTGCTTGCA**TAA**CTGATCATATTGCTTGTCTCAGAAAAAACACCATCAGCAAAG  
TGGGCCTCCAAAGCCAGATGATTTTCATTTTATGTGTGAATAATAGCTTCATTAACACAATCA  
AGACCATGCACATTGTAAATACATTATTCTTGGATAATTCTATACATAAAAGTTCCTACTTGT  
TAAA



446/615

**FIGURE 442**

MAVKLGTLALLALGLAQPASARRKLLVFLLDGFRSDYISDEALES LPGFKEIVSRGVKVDYL  
TPDFPSLSYPNYYTLM TGRHCEVHQMIGNYMWDPTTNKSF DIGV NKDSL MPLWWNGSEPLWVT  
LTKAKRKVYMYYPGCEVEILGVRPTYCLEYKNVPTDINFANAVSDALDSFKSGRADLA AIYH  
ERIDVEGHHYGPASPQRKDALKAVDTVLKYMTKWIQERGLQDRLNVIIFSDHGMTDIFWMDKV  
IELNKYISLNDLQQVKDRGPVVS LWPAPGKHSEIYNKLSTVEHMTVYEKEAIPSRFYKKGKF  
VSPLTLVADEGWFITENREMLPFWMNSTGRREGWQRGWHGYDNELMDMRGIFLAFGPDFKS NF  
RAAPIRSVDVYNVMCNVVGITPLPNNGSWSRVMCMLKGRAGTAPPVWPSHCALALILLFLLA

**Important features of the protein:****Signal peptide:**

amino acids 1-22

**N-glycosylation sites.**

amino acids 100-104, 118-122, 341-345, 404-408

**N-myristoylation sites.**

amino acids 148-154, 365-371

**Amidation site.**

amino acids 343-347

447/615

**FIGURE 443**

AGTGACTGCAGCCTTCCTAGATCCCCTCCACTCGGTTTCTCTCTTTGCAGGAGCACCGGCAGC  
ACCA GTGTGTGAGGGGAGCAGGCAGCGGTCTAGCCAGTTCCTTGATCCTGCCAGACCACCCA  
GCCCCCGGCACAGAGCTGCTCCACAGGCACCA**ATG**AGGATCATGCTGCTATTACAGCCATCCT  
GGCCTTCAGCCTAGCTCAGAGCTTTGGGGCTGTCTGTAAGGAGCCACAGGAGGAGGTGGTTCC  
TGGCGGGGGCCGCAGCAAGAGGGATCCAGATCTCTACCAGCTGCTCCAGAGACTCTTCAAAG  
CCACTCATCTCTGGAGGGATTGCTCAAAGCCCTGAGCCAGGCTAGCACAGATCCTAAGGAATC  
AACATCTCCCGAGAAACGTGACATGCATGACTTCTTTGTGGGACTTATGGGCAAGAGGAGCGT  
CCAGCCAGAGGGAAAGACAGGACCTTTCTTACCTTCAGTGAGGGTTCCTCGGCCCTTCATCC  
CAATCAGCTTGATCCACAGGAAAGTCTTCCCTGGGAACAGAGGAGCAGAGACCTTTA**TAAGA**  
CTCTCCTACGGATGTGAATCAAGAGAACGTCCCCAGCTTTGGCATCCTCAAGTATCCCCCGAG  
AGCAGAATAGGTACTCCACTTCCGGACTCCTGGACTGCATTAGGAAGACCTCTTTCCCTGTCC  
CAATCCCCAGGTGCGCACGCTCCTGTTACCTTTCTCTTCCCTGTTCTTGTAACATTCTTG  
CTTTGACTCCTTCTCCATCTTTTCTACCTGACCCTGGTGTGGAACTGCATAGTGAATATCCC  
CAACCCCAATGGGCATTGACTGTAGAATAACCTAGAGTTCCTGTAGTGTCTACATTAAAAAT  
ATAATGTCTCTCTCTATTCCTCAACAATAAAGGATTTTGCATATGAAAAAAAAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAA

448/615.

**FIGURE 444**

MRIMLLFTAILAFSLAQSFQAVCKEPQEEVVPGGGRSKRDPDLYQLLQRLFKSHSSLEGLLKA  
LSQASTDPKESTSPEKRDMDHFFVGLMGKRSVQPEGKTGPFLPSVRVPRPLHPNQLGSTGKSS  
LGTEEQRPL

**Important features:****Signal peptide:**

amino acids 1-18

**Tyrosine kinase phosphorylation site.**

amino acids 36-45

**N-myristoylation site.**

amino acids 33-39, 59-65

amino acids 43-65

**Tachykinin family signature.**

amino acids 86-92

449/615

**FIGURE 445**

TGGACTTCTCTGGACCACAGTCCTCTGCCAGACCCCTGCCAGACCCAGTCCACCATGATCCATCTGGGTCACAT  
CCTCTTCCTGCTTTTGTCTCCAGTGGCTGCAGCTCAGACGACTCCAGGAGAGAGATCATCACTCCCTGCCTTTTA  
CCCTGGCACTTCAGGCTCTTGTTCCGGATGTGGGTCCCTCTCTCTGCCGCTCCTGGCAGGCCTCGTGGCTGCTGA  
TGCGGTGGCATCGCTGCTCATCGTGGGGGCGGTGTTCTGTGCGCACGCCCACGCCGCAGCCCCGCCCAAGATGG  
CAAAGTCTACATCAACATGCCAGGCAGGGGCTGACCCCTCCTGCAGCTTGGACCTTTGACTTCTGACCCTCTCATC  
CTGGATGGTGTGTGGTGGCACAGGAACCCCGCCCCAACTTTGGATTGTAATAAAACAATTGAAACACCA

450/615

**FIGURE 446**

MIHLGHILFLLLLPVAAAQTTPGERSSSLPAFYPTSGSCSGCSLSLPLLAGLVAADAVASLLIVGAVFLCARPR  
RSPAQDGKVYINMPGRG

Signal peptide:	Amino acids 1-18
transmembrane domain:	Amino acids 51-70
Glycosaminoglycan attachment site:	Amino acids 40-44
N-myristoylation sites:	Amino acids 34-40;37-43;52-58
Prokaryotic membrane lipoprotein lipid attachment site:	Amino acids 29-40

451/615

**FIGURE 447**

GCCAGGTGTGCAGGCCGCTCCAAGCCCAGCCTGCCCCGCTGCCGCCACCATGACGCTCCTCCC  
CGGCCTCCTGTTTCTGACCTGGCTGCACACATGCCTGGCCACCATGACCCCTCCCTCAGGGG  
GCACCCCCACAGTCACGGTACCCACACTGCTACTCGGCTGAGGAAGTCCCCCTCGGCCAGGC  
CCCCCCACACCTGCTGGCTCGAGGTGCCAAGTGGGGGCAGGCTTTCCTGTAGCCCTGGTGTC  
CAGCCTGGAGGCAGCAAGCCACAGGGGGAGGCACGAGAGGCCCTCAGCTACGACCCAGTGCCC  
GGTGCTGCGGCCGGAGGAGGTGTTGGAGGCAGACACCCACCAGCGCTCCATCTCACCCCTGGAG  
ATACCGTGTGGACACGGATGAGGACCGCTATCCACAGAAGCTGGCCTTCGCCGAGTGCCTGTG  
CAGAGGCTGTATCGATGCACGGACGGGCCGCGAGACAGCTGCGCTCAACTCCGTGCGGCTGCT  
CCAGAGCCTGCTGGTGCTGCGCCGCCGCCCTGCTCCCGCGACGGCTCGGGGCTCCCCACACC  
TGGGGCCTTTGCCTTCCACACCGAGTTCATCCACGTCCCCGTCGGCTGCACCTGCGTGCTGCC  
CCGTTCAGTGTTGACCGCCGAGGCCGTGGGGCCCCCTAGACTGGACACGTGTGCTCCCCAGAGGG  
CACCCCTATTTATGTGTATTTATTGTTATTTATATGCCTCCCCAACACTACCCTTGGGGTC  
TGGGCATTCCCCGTGTCTGGAGGACAGCCCCCACTGTTCTCCTCATCTCCAGCCTCAGTAGT  
TGGGGGTAGAAGGAGCTCAGCACCTCTTCCAGCCCTTAAAGCTGCAGAAAAGGTGTCACACGG  
CTGCCTGTACCTTGGCTCCCTGTCCTGCTCCCGGCTTCCCTTACCCTATCACTGGCCTCAGGC  
CCCGCAGGCTGCCTCTTCCCAACCTCCTTGGAAGTACCCCTGTTTCTTAAACAATTATTTAAG  
TGTACGTGTATTATTAAACTGATGAACACATCCCCAAA

452/615

## **FIGURE 448**

MTLLPGLLFLTWLHTCLAHHDPSLRGHPHSHGTPHCYSAEELPLGQAPPHLLARGAKWGQALP  
VALVSSLEAASHRGRHERPSATTQCPVLRPEEVLEADTHQRSISPWRYRVDTDEDYPQKLAF  
AECLCRGCIDARTGRETAALNSVRLQSLVLRRRPCSRDGSGGLPTPGAFAFHTEFIHVPVGC  
TCVLPRSV

**Important features:**

**Signal peptide:**

amino acids 1-18

**Tyrosine kinase phosphorylation site.**

amino acids 112-121

**N-myristoylation sites.**

amino acids 32-38, 55-61, 133-139

amino acids 3-25

**Homologous region to IL-17.**

amino acids 99-195

453/615

**FIGURE 449**

TGCAGAGCTTGTGGAGGCC**ATG**GGGCGCGTCGTCGCGGAGCTCGTCTCCTCGCTGCTGGGGTT  
GTGGCTGTTGCTGTGCAGCTGCGGATGCCCCGAGGGCGCCGAGCTGCGTGCTCCGCCAGATAA  
AATCGCGATTATTGGAGCCGGAATTGGTGACACTTCAGCAGCCTATTACCTGCGGCAGAAATT  
TGGGAAAGATGTGAAGATAGACCTGTTTGAAAGAGAAGAGGTCGGGGGCCGCTGGCTACCAT  
GATGGTGCAGGGGCAAGAATACGAGGCAGGAGGTTCTGTCTATCCATCCTTTAAATCTGCACAT  
GAAACGTTTTGTCAAAGACCTGGGTCTCTCTGCTGTTTCAGGCCTCTGGTGGCCTACTGGGGAT  
ATATAATGGAGAGACTCTGGTATTTGAGGAGAGCAACTGGTTCATAATTAACGTGATTAAATT  
AGTTTGGCGCTATGGATTTCAATCCCTCCGTATGCACATGTGGGTAGAGGACGTGTTAGACAA  
GTTTCATGAGGATCTACCGCTACCACTCTCATGACTATGCCTTCAGTAGTGTCGAAAAATTACT  
TCATGCTCTAGGAGGAGATGACTTCCTTGGAATGCTTAATCGAACACTTCTTGAAACCTTGCA  
AAAGGCCGGCTTTTCTGAGAAGTTCCTCAATGAAATGATTGCTCCTGTTATGAGGGTCAATTA  
TGGCCAAAGCACGGACATCAATGCCTTTGTGGGGCGGTGTCACTGTCCTGTTCTGATTCTGG  
CCTTTGGGCAGTAGAAGGTGGCAATAAACTTGTTTGCTCAGGGCTTCTGCAGGCATCCAAAAG  
CAATCTTATATCTGGCTCAGTAATGTACATCGAGGAGAAAACAAAGACCAAGTACACAGGAAA  
TCCAACAAAGATGTATGAAGTGGTCTACCAAATTGGAAGTGAAGTTCGTTTCAGACTTCTATGA  
CATCGTCTTGGTGGCCACTCCGTTGAATCGAAAAATGTCGAATATTACTTTTCTCAACTTTGA  
TCCTCCAATTGAGGAATTCCATCAATATTATCAACATATAGTGACAACTTTAGTTAAGGGGGA  
ATTGAATACATCTATCTTTAGCTCTAGACCCATAGATAAATTTGGCCTTAATACAGTTTTAAC  
CACTGATAATTCAGATTTGTTTCATTAAACAGTATTGGGATTGTGCCCTCTGTGAGAGAAAAGGA  
AGATCCTGAGCCATCAACAGATGGAACATATGTTTGGAAGATCTTTTCCCAAGAACTCTTAC  
TAAAGCACAAATTTTAAAGCTCTTTCTGTCTTATGATTATGCTGTGAAGAAGCCATGGCTTGC  
ATATCCTCACTATAAGCCCCGGAGAAATGCCCTCTATCATTCTCCATGATCGACTTTATTA  
CCTCAATGGCATAGAGTGTGCAGCAAGTGCCATGGAGATGAGTGCCATTGCAGCCACAACGC  
TGCACTCCTTGCCATCACCGCTGGAACGGGCACACAGACATGATTGATCAGGATGGCTTATA  
TGAGAACTTAAACTGAAGT**TGA**AGTGACACACTCCTTTTTCCCTCCTAGTTCCAAATGA  
CTATCAGTGGCAAAAAAGAACAAAATCTGAGCAGAGATGATTTTGAACCAGATATTTTGCCAT  
TATCATTGTTTAATAAAAGTAATCCCTGCTGGTCATAGGAAAAAAAAAAAAA



454/615

**FIGURE 450**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA62880
<subunit 1 of 1, 505 aa, 1 stop
<MW: 56640, pI: 6.10, NX(S/T): 4
MGRVVAELVSSLLGLWLLLCSCGCPEGAELRAPDPKIAIIGAGIGGTSAAYYLRQKFGKDVKI
DLFEREEVGGRLATMMVQGEYEAGGSVIHPLNLHMKRFVKDLGLSAVQASGGLLGIYNGETL
VFEESNWFIIINVIKLVWRYGFQSLRMHMMWVEDVLDKFMRIYRYQSHDYAFSSVEKLLHALGGD
DFLGMLNRTLLETLOKAGFSEKFLNEMIAFVMRVNYGQSTDINAFVGAVSLSCSDSGLWAVEG
GNKLVCSEGLLQASKSNLISGSVMYIEEKTCTKYTGNTKMYEVVYQIGTETRSDFYDIVLVAT
PLNRKMSNITFLNFDPPIEEFHQYYQHIVTTLVKGELNTSIFSSRPIDKFGINTVLTTDNSDL
FINSIGIVPSVREKEDPEPSTDGTYVWKIFSQETLTQAQILKLFLSYDYAVKKPWLAYPHYKP
PEKCPSIILHDLRYLLNGIECAASAMEMSAIAAHNAALLAYHRWNGHTDMIDQDGLYEKLKTEL
```

**Important features:****Signal peptide:**

amino acids 1-23

**N-glycosylation sites.**

amino acids 196-200, 323-327, 353-357

**Tyrosine kinase phosphorylation site.**

amino acids 291-298

**N-myristoylation sites.**amino acids 23-29, 41-47, 43-49, 45-51, 46-52, 72-78, 115-121,  
119-125, 260-266, 384-390, 459-465**Prokaryotic membrane lipoprotein lipid attachment site.**

amino acids 12-23, 232-243

455/615

**FIGURE 451**

[illegible]

456/615

**FIGURE 452**

MASYLYGVLEFAVGLCAPIYCVSPANAPSAYPRPSSTKSTPASQVYSLNTDEAFRLYRRLVLET  
PSQNIFFSPVSVSTSLAMLSLGAHSVTKTQILQGLGFNLTHTPESAIHQGFQHLVHSLTVPSK  
DLTLKMGSALFVKKELQLQANFLGNVKRLYEAEVFSTDFSNPSIAQARINSHVKKKTQGKVVD  
IIQGLDLLTAMVLVNHIFFKAKWEKPFHLEYTRKNFPFLVGEQVTVQVPMMHQKEQFAFGVDT  
ELNCFVLQMDYKGDVAFFVLPSKGMROLEQALSARTLIKWSHSLQKRWIEVFIPRFSISAS  
YNLETILPKMGIQNAFDKNADFSGIAKRDSLQVSKATHKAVLDVSEEGTEATAATTTKFIVRS  
KDGPSYFTVSFNRTFLMMITNKATDGILFLGKVENPTKS

**Signal peptide:**  
amino acids 1-20

457/615

**FIGURE 453**

CTCCGGGTCCCCAGGGGCTGCGCCGGGCCGGCCTGGCAAGGGGGACGAGTCAGTGGACACTCCAGGAAGAGCGGC  
CCCCGGGGGGGCGATGACCGTGCGCTGACCCTGACTCACTCCAGGTCCGGAGGCGGGGGCCCCGGGGCGACTCG  
GGGGCGGACCGCGGGGGCGGAGCTGCCGCCCGTGAGTCCGGCCGAGCCACCTGAGCCCCGAGCCGCGGGACACCGTC  
GCTCCTGCTCTCCGAATGCTGCGCACCGCGATGGGCCTGAGGAGCTGGCTCGCCGCCCATGGGGCGCGCTGCCG  
CCTCGGCCACCGCTGCTGCTGCTCCTGCTGCTGCTGCTCCTGCTGCAGCCGCCGCTCCGACCTGGGCGCTCAGC  
CCCCGGATCAGCCTGCCCTCTGGGCTCTGAAGAGCGGGCCATTCCTCAGATTCTGAAGCTGAACACATCTCCAACCTAC  
ACAGCCCTTCTGCTGAGCAGGGATGGCAGGACCCTGTACGTGGGTGCTCGAGAGGCCCTCTTTGCACTCAGTAGC  
AACCTCAGCTTCTGCCAGGCGGGGAGTACCAGGAGCTGCTTTGGGGTGCAGACGCAGAGAAGAAACAGCAGTGC  
AGCTTCAAGGGCAAGGACCCACAGCGGAGTGTCAAACCTACATCAAGATCCTCCTGCCGCTCAGCGGCAGTCAC  
CTGTTACCTGTGGCACAGCAGCCTTCAGCCCCATGTGTACCTACATCAACATGGAGAATTCACCTTGGCAAGG  
GACGAGAAGGGGAATGTCTCCTGGAAGATGGCAAGGGCCGTTGTCCCTTCGACCCGAATTTCAAGTCCACTGCC  
CTGGTGGTTGATGGCGAGCTCTACACTGGAACAGTCAGCAGCTTCCAAGGGAATGACCCGGCCATCTCGCGGAGC  
CAAAGCCTTCGCCCCACCAAGACCGAGAGCTCCCTCAACTGGCTGCAAGACCCAGCTTTTGTGGCCTCAGCCTAC  
ATTCTGAGAGCCTGGGCAGCTTGCAAGGCGATGATGACAAGATCTACTTTTCTTCAGCGAGACTGGCCAGGAA  
TTTGAGTTCTTTGAGAACACCATTTGTGTCCCGCATTGCCCGCATCTGCAAGGGCGATGAGGGTGGAGAGCGGGTG  
CTACAGCAGCGCTGGACCTCCTTCCCTCAAGGCCAGCTGCTGTGCTCACGGCCCGACGATGGCTTCCCCTTCAAC  
GTGCTGCAGGATGTCTTCACGCTGAGCCCCAGCCCCAGGACTGGCGTGACACCTTTTCTATGGGGTCTTCACT  
TCCCAGTGGCACAGGGGAACCTACAGAAGGCTCTGCGCTGTGTCTTCAATGAAGGATGTGCAGAGAGTCTTC  
AGCGGCTCTACAAGGAGGTGAACCGTGAGACACAGCAGTGGTACACCGTGACCCACCCGGTGCCACACCCCGG  
CCTGGAGCGTGATCACCAACAGTGCCCCGGGAAGGAAGATCAACTCATCCCTGCAGCTCCAGACCCGCTGCTG  
AACTTCTCAAGGACCACTTCTGATGGACGGGCAGTCCGAAGCCGATGCTGCTGCTGCAGCCCCAGGCTCGC  
TACCAGCGCGTGGCTGTACACCGCGTCCCTGGCCTGCACCACACCTACGATGTCTCTTCTTGGGCACTGGTGAC  
GGCCGGCTCCACAAGGCAGTGAGCGTGGGCCCCGGGTGCACATCATTTGAGGAGCTGCAGATCTTCTCATCGGGA  
CAGCCCGTGCAAGATCTGCTCCTGGACCCACAGGGGGCTGCTGTATGCGGCCCTCACACTCGGGCGTAGTCCAG  
GTGCCATGGCCAACCTGCAGCCTGTACCGGAGCTGTGGGACTGCCTCCTCGCCCCGGGACCCCTACTGTGCTTGG  
AGCGGCTCCAGCTGCAAGCAGCTCAGCCTCTACCAGCCTCAGCTGGCCACCAGGCGGTGGATCCAGGACATCGAG  
GGAGCCAGCGCCAAGGACCTTTGCAGCGCGTCTTCGGTTGTGTCCCCGTCTTTTGTACCAACAGGGGAGAAGCCA  
TGTGAGCAAGTCCAGTTCAGCCCCACACAGTGAACACTTTGGCCTGCCCGCTCCTCTCCAACCTGGCGACCCGA  
CTCTGGCTACGCAACGGGGCCCCCGTCAATGCCTCGGCCTCCTGCCACGTGCTACCCACTGGGGACCTGCTGCTG  
TGGGGACCCCAACAGCTGGGGGAGTTCCAGTGCTGGTCACTAGAGGAGGGCTTCCAGCAGTGGTAGCCAGCTAC  
TGCCAGAGGTGGTGGAGACGGGGTGGCAGACCAACAGATGAGGGTGGCAGTGTACCCGTCAATTATCAGCACA  
TCGCGTGTGAGTGACACAGCTGGTGGCAAGGCCAGCTGGGGTGCAGACAGGTCTACTGGAAGGAGTTCTGGTG  
ATGTGCACGCTCTTTGTGCTGGCCGTGCTGCTCCAGTTTATTTCTTGTCTTACCGGCACCGGAACAGCATGAAA  
GTCTTCTGAAGCAGGGGGAATGTGCCAGCGTGACCCCAAGACCTGCCCTGTGGTGCTGCCCCCTGAGACCCGC  
CCACTCAACGGCCTAGGGCCCCCTAGCACCCCGCTCGATCACCGAGGGTACCAGTCCCTGTGAGACAGCCCCCG  
GGGGCCCCGAGTCTTCACTGAGTCAGAGAAGAGGCCACTCAGCATCCAAGACAGCTTCGTGGAGGTATCCCCAGTG  
TGCCCCCGGGCCCCGGTCCGCCTTGGCTCGGAGATCCGTGACTCTGTGGTGTGAGAGCTGACTTCCAGAGGACGC  
TGCCCTGGCTTCAGGGGCTGTGAATGCTCGGAGAGGGTCAACTGGACCTCCCTCCGCTCTGCTCTTCTGGAAC  
ACGACCGTGGTGCCCGGCCCTTGGGAGCCTTGGAGCCAGCTGGCCTGCTGCTCTCCAGTCAAGTAGCGAAGCTCC  
TACCACCCAGACACCCAAACAGCCGTGCCCCAGAGGTCTGGCCAAATATGGGGGCTGCCTAGGTGGTGGA  
CAGTGCTCCTTATGTAACTGAGCCCTTTGTTTAAAAAACAATTCCAAATGTGAACTAGAATGAGAGGGAAGAG  
ATAGCATGGCATGCAGCACACAGGCTGCTCCAGTTTATGGCCTCCAGGGGTGCTGGGGATGCATCCAAAGTGG  
TTGTCTGAGACAGAGTTGGAAACCTCACCAACTGGCCTTTCACCTTCCACATTATCCCGCTGCCACCGGCTGC  
CCTGTCTCACTGCAGATTACAGGACAGCTTGGGCTGCGTGCGTTCTGCCTTGCCAGTCAGCCGAGGATGTAGTTG  
TTGCTGCCGTGCTCCACCACCTCAGGGACAGAGGGCTAGGTTGGCACTGCGGCCCTCACCAGTCTGGGCTC  
GGACCCAACTCCTGGACCTTTCCAGCCTGTATCAGGCTGTGGCCACACGAGAGGACAGCGCGAGCTCAGGAGAGA  
TTTCGTGACAATGTACGCCTTTCCCTCAGAATTCAGGGAAGAGACTGTGCGCTGCCTTCCCTCCGTTGTTGCGTGA  
GAACCCGTGTGCCCTTTCCACCATATCCACCCTCGCTCCATCTTTGAACTCAAACACGAGGAACCTAAGTGCACC  
CTGGTCTTCCCGAGTCCCGAGTTCACCTCCATCCCTCCACTTCCCTCCACTCTAAGGGATATCAACACTGCCC  
AGCACAGGGGCCCTGAATTTATGTGGTTTTTATACATTTTTTAATAAGATGCACCTTATGTCATTTTTTAATAAA  
GTCTGAAGAATTACTGTTTAAAAAATAA

458/615

**FIGURE 454**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA67962
><subunit 1 of 1, 837 aa, 1 stop
><MW: 92750, pI: 7.04, NX(S/T): 6
MLRTAMGLRSWLAAPWGALPPRPPLLLLLLLLLLLLLLQPPPPTWALSPRISLPLGSEERPFLRFE
AEHISNYTALLLSRDGRTLYVGAREALFALSSNLSFLPGGEYQELLWGADAEKKQQCSFKGKD
PQRDCQNYIKILLPLSGSHLFTCGTAAFSPMCTYINMENFTLARDEKGNVLLLEDGKGRCPFDP
NEKSTALVVDGELYTGTVSSFQGNDDPAISRSQSLRPTKTESSLNWLQDPAFVASAYIPESLGS
LQGDDDKIYFFFSETGQEFEEFFENTIVSRIARICKGDEGGERVLQQRWTSFLKAQLLCSRDD
GFPFENVLQDVFTLSPPQDWRDTLFGVFTSQWHRGTTEGSAVCVFTMKDVQRVFSGLYKEVN
RETQQWYTVTHPVPTPRPGACITNSARERKINSSLQLPDRVLNFKDHFLMDGQVRSRMLLLQ
PQARYQRVAVHRVPGLHHTYDVLFLGTGDGRLHKAVSVGPRVHIIIEELQIFSSGQPVQNLLLD
THRGLLYAASHSGVVQVPMANCSLYRSCGDCLLARDPYCAWSGSSCKHVSPLYQPQLATRPWIQ
DIEGASAKDLCSASSVSPSFVPTGEKPCQVQFQPNVTNLTACPLLSNLATRLWLRNGAPVN
ASASCHVLPTGDLLLVTQQLGEFQCWSLEEGFQQLVASYCPEVVEDGVADQTDEGGSPVPII
STSRVSAPAGGKASWGADRSYWKEFLVMCTLFVLAVLLPVLFLLYRHRNSMKVFLKQGECAV
HPKTCPVVLPPEPTRPLNGLGPPSTPLDHRGYQSLSDSPPGARVFTSEKRPLSIQDSFVEVSP
VCPRPRVRLGSEIRDSV
```

**Transmembrane domains:**

amino acids 23-46 (type II), 718-738

459/615

**FIGURE 455**

TAAGATGAGGGCATCCCTCACGTTACACCCCCTGGTGGCATCTGCCAGCCCTGTTCTGGGGAC  
AAGGCGGGCTTTTCGTGGGAGCCATGCTCAGCCTGCCAGGAAGCCAAGCCCTACAGTGCAGAGG  
AAACAGAATTTCAACGGGAAGCTGGTTTGCTTCATACCATTGGGATCTGCTGGTAAAGCTGTT  
ATTTGGGTTTAGGGACTGATCCCTTGACGTTTACTTCTGGATCACCATGAATGGCCAAGATGG  
TGGCAGAACACGCTGTGGACCCCTGAGTTAGAGACAATGCAAATGTTGGATTGGGTGTAATTCT  
TTTTGAATCCCAGATCCAGTCTGTACTTGAATATGAGCAGAAGATCTACAAGAATGCTGACAG  
GGAACCGTGTTAAGACCCAGCACCCCTATTCCCAGGAGCTTCTGGCCTGACCATCTGCAGCCA  
AAGCACTAACAGGGACAGATATGGGAATGTCCACCTTTGATCCGCATCCTGCACAATAGTGGT  
CCCACCATGGCTGCCACTTTTTTATACTATTTGGAGAAAAGACCTTGTATAAATTCGAGGCCC  
GAGTGAATAACGTCTCTGTACACCGAAATGGGTACTTGGTGGCATAGAGAAACACAATTAGC  
CACTTTTTTCAGCTACACTTCTCACTCAGCTGCACCCCTACACTTCTCACTCAGGTGCACCCCT  
TCTGCTGTCTTTTCCCAACGTACTGGGTCCCAGCGTGGTGGGTATTTGCCACACTGGGTGC  
CAGCTCAGCAGCCCCCACCCTCTCTTTATTCTCTCCAAAGCTGGTCTTTCTGACTATCATTGT  
GGTAGGGGGAGGACAGATGCTAAAGGTGGAAGCTGACCTGGAGAAAGAGACACACGGGGTGAC  
TGTGGCAAAGGACAGCTGGAAAAGAACTCTATCACTTCTTCATTGGCAACCACAAGGCACCC  
GAGGCCATGGCACTCCCAGAGGCTGTGCGCAGAGCCAAGCCTCTCAACCTCTTCTGGCCCTGC  
GTCCTGCAGCGAAGTCTCTGCTGTAAGACAGTAGACTCCTTCGATGAGGTGCTCAAAAATGCT  
ACCCGGGGTGGTGGTGTCTGGCTTGCAGTCTGGCCAGTTCAGAGAAAGTTGCAGAGATCAGGG  
GCCAAGGATGTCATAGCCCCAGGTTGTCTCAGGGTCCAATCCTAGGGCAGGGTGTGCATGG  
AAGCAAGAACTATGGAAACCTAGCTCCAGTCTGCAGGCTCTGAGCCCCCTAGTTCTCTCACTCCA  
GCGGGGCTCCCTCACTGCACAGAACCCACCCCTTCTGTGTGGGCACTGCTGACCACACAGATG  
ACCCAGACCCAAAGAGCCTGGCAGAAGCTCTGTGGTTGGAGCTGGGCTCCGTCTCCAGGTCTG  
GTTCAGGGGGATCAGGAAGGCTCTTTTCCACCTGTGGCTTCACTGGCCCTTTGAGATTTCTTA  
TCTCACCGTTACTTCAGTTACCCTTGCAGGGGGCCAGGGAGTCAAGAATATACCGTGTTCCTC  
CAGGGTTTAAGCCGGCCATGCCTTCCCCGAGAGCATAACCAACTTGACAGGGGTGCCAGTTAC  
CCCACAACTGAAGGAAGGAGATCCTTCCCCCGTCCCAGGAGTGCTCTCAACCAGCCTCAGA  
AAGCTTGAGAAGATGGACCCTTTGCCCACCAGGGTTAATTCTTGGTGGGGCAGCTCGGCTGTG  
ATCAGGGCAACCAAACCTATAGGAAGCCTTCCAGTGTGAGCTGGAATTAGACTGAACATGTGC  
TTGGGCCTGCCTCTCCCTAGACGCAGTTGCGGGGCACTCCAGGGAATGAACCAGCTCAAGTGT  
GTCCTAACAGCAGCCTGGAGCTACCCCCAATCCCTCACAGCCTGACCCTCTCATTCCATCA  
GATCTCGTGCCG

460/615

**FIGURE 456**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA69555  
><subunit 1 of 1, 148 aa, 1 stop  
><MW: 16214, pI: 10.22, NX(S/T): 0

MGTWWHRETQLATFSATLLTQLHPTLLTQVHPLLLSFPQRTGSRAWVVFATLGASSAAPH  
LSLFSPKLVFLTIIVVGGGQMLKVEADLEKETHGVTVAKDSWKRNSITSSLATTRHPRPW  
HSQRLCAEPSLSTSSGPASCSEVSAVRQ

**Important features of the protein:****Signal peptide:**

Amino acids 1-28

**Transmembrane domain:**

Amino acids 64-78

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 103-107

**N-myristoylation sites:**

Amino acids 53-59;94-100

461/615

**FIGURE 457**

CCCGCGCGCCCCTGGCACTCAATCCCCGCC**ATG**TGGGGGCTCCTGCTCGCCCTGGCCGCCTTC  
GCGCCGCGCGTTCGGCCCGGCTCTGGGGGCGCCAGGAACTCGGTGCTGGGCCTCGCGCAGCCC  
GGGACCACCAAGGTCCCAGGCTCGACCCCGGCCCTGCATAGCAGCCCGGCACAGCCGCCGGCG  
GAGACAGCTAACGGGACCTCAGAACAGCATGTCCGGATTTCGAGTCATCAAGAAGAAAAAGGTC  
ATTATGAAGAAGCGGAAGAAGCTAACTCTAACTCGCCCCACCCCACTGGTGACTGCCGGGCCC  
CTTGTGACCCCCACTCCAGCAGGGACCCCTCGACCCCGCTGAGAAACAAGAAACAGGCTGTCTT  
CCTTTGGGTCTGGAGTCCCTGCGAGTTTCAGATAGCCGGCTTGAGGCATCCAGCAGCCAGTCC  
TTTGGTCTTGGAACACACCGAGGACGGCTCAACATTCATTTCAGGCCTGGAGGACGGCGATCTA  
TATGATGGAGCCTGGTGTGCTGAGGAGCAGGACGCCGATCCATGGTTTCAGGTGGACGCTGGG  
CACCCACCCGCTTCTCGGGTGTTATCACACAGGGCAGGAACTCTGTCTGGAGGTATGACTGG  
GTCACATCATAACAAGGTCCAGTTCAGCAATGACAGTCGGACCTGGTGGGGAAGTAGGAACCAC  
AGCAGTGGGATGGACGCAGTATTTCTGCCAATTCAGACCCAGAACTCCAGTGCTGAACCTC  
CTGCCGGAGCCCCAGGTGGCCCGCTTCATTCGCCTGCTGCCCCAGACCTGGCTCCAGGGAGGC  
GCGCCTTGCTCCGGGCAGAGATCCTGGCCTGCCAGTCTCAGACCCCAATGACCTATTCTCTT  
GAGGCCCTGCGTCGGGATCCTCTGACCCTCTAGACTTTTCAGCATCACAATTACAAGGCCATG  
AGGAAGCTGATGAAGCAGGTACAAGAGCAATGCCCCAACATCACCCGCATCTACAGCATTGGG  
AAGAGCTACCAGGGCCTGAAGCTGTATGTGATGGAAATGTCGGACAAGCCTGGGGAGCATGAG  
CTGGGGGAGCCTGAGGTGCGCTACGTGGCTGGCATGCATGGGAACGAGGCCCTGGGGCGGGAG  
TTGCTTCTGCTCCTGATGCAGTTCCTGTGCCATGAGTTCCTGCGAGGGAACCCACGGGTGACC  
CGGCTGCTCTCTGAGATGCGCATTACCTGCTGCCCTCCATGAACCCTGATGGCTATGAGATC  
GCCTACCACCGGGGTTTCAGAGCTGGTGGGCTGGGCCGAGGGCCGCTGGAACAACCAGAGCATC  
GATCTTAACCATAATTTTGCTGACCTCAACACACCACTGTGGGAAGCACAGGACGATGGGAAG  
GTGCCCCACATCGTCCCCAACCATCACCTGCCATTGCCCACTTACTACACCCTGCCCAATGCC  
ACCGTGGCTCCTGAAACGCGGGCAGTAATCAAGTGGATGAAGCGGATCCCCCTTTGTGCTAAGT  
GCCAACCTCCACGGGGGTGAGCTCGTGGTGTCTTACCCATTTCGACATGACTCGCACCCCGTGG  
GCTGCCCCGCGAGCTCACGCCCACACCAGATGATGCTGTGTTTCGCTGGCTCAGCACTGTCTAT  
GCTGGCAGTAATCTGGCCATGCAGGACACCAGCCGCCGACCCTGCCACAGCCAGGACTTCTCC  
GTGCACGGCAACATCATCAACGGGGCTGACTGGCACACGGTCCCCGGGAGCATGAATGACTTC  
AGCTACCTACACACCAACTGCTTTGAGGTCAGTGTGGAGCTGTCCTGTGACAAGTTCCCTCAC  
GAGAATGAATTGCCCCAGGAGTGGGAGAACAACAAAGACGCCCTCCTCACCTACCTGGAGCAG  
GTGCGCATGGGCATTGCAGGAGTGGTGGGGACAAGGACACGGAGCTTGGGATTGCTGACGCT  
GTCATTGCCGTGGATGGGATTAACCATGACGTGACCACGGCGTGGGGCGGGGATTATTGGCGT  
CTGCTGACCCCAAGGGACTACATGGTGAAGTGGCAGTGGCGAGGGCTACCATTTCAGTGACACGG  
AACTGTCCGGTACCTTTGAAGAGGGCCCCCTTCCCCCTGCAATTTTCGTGCTCACCAAGACTCCC  
AAACAGAGGCTGCGCGAGCTGCTGGCAGCTGGGGCCAAGGTGCCCCCGGACCTTCGCAGGCGC  
CTGGAGCGGCTAAGGGGACAGAAGGAT**TGA**TACCTGCGGTTTAAGAGCCCTAGGGCAGGCTGG  
ACCTGTCAAGACGGGAAGGGGAAGAGTAGAGAGGGAGGGACAAAGTGAGGAAAAGGTGCTCAT  
TAAAGCTACCGGGCACCTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA



462/615

**FIGURE 458**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71162
><subunit 1 of 1, 734 aa, 1 stop
><MW: 81677, pI: 6.60; NX(S/T): 6
MWGLLLALAAFAPAVGPALGAPRNSVLGLAQPGETTKVPGSTPALHSSPAQPPAETANGTS
EQHVRIRVIKKKKVIMKKRKKLTLTRPTPLVTAGPLVTPTPAGTLDPAEKQETGCPPLGL
ESLRVSDSRLEASSSSQSFGGLGPHRGRNLNIHSGLEDGDLYDGAWCAEEQDADPWFQVDAGH
PTRFSGVITQGRNSVWRYDWVTSYKVQFSNDSRTWWGSRNHSSGMDAVFPANSDPETPVL
NLLPEPQVARFIRLLPQTWLQGGAPCLRAEILACPVSDPNDLFLEAPASGSSDPLDFQHH
NYKAMRKLMKQVQEQCPNITRIYSIGKSYQGLKLYVMEMSDKPGEHELGEPEVRYVAGMH
GNEALGRELLLLLMQFLCHEFLRGNPRVTRLLSEMRIHLLPSMNPDGYEIAHYHRGSELVG
WAEGRWNNQSIDLNNHFADLNTPLWEAQDDGKVP HIVPNHHLPLPTYYTLPNATVAPETR
AVIKWMKRIPFVLSANLHGGELVVSYPFDMTRTPWAARELTPTPDDAVFRWLSTVYAGSN
LAMQDTSRRPCHSQDFSVHGNIINGADWHTVPGSMNDFSYLHTNCFEVTVELSCDKFPHE
NELPQEWENNKDALLTYLEQVRMGIAGVVRDKDTELGIADAVIAVDGINHDVTTAWGGDY
WRLLTTPGDYMTASAEGYHSVTRNCRVTFEEGPFPCNFVLTKTPKQRLRELLAAGAKVPP
DLRRRLERLRGQKD
```

463/615

**FIGURE 459**

TAAAACAGCTACAATATTCCAGGGCCAGTCACTTGCCATTTCTCATAACAGCGTCAGAGAGAA  
AGAACTGACTGAAACGTTTGAGATGAAAGAAAGTTCTCCTCCTGATCACAGCCATCTTGGCAGT  
GGCTGTTGGTTTCCCAGTCTCTCAAGACCAGGAACGAGAAAAAAGAAGTATCAGTGACAGCGA  
TGAATTAGCTTCAGGGTTTTTTGTGTTCCCTTACCCATATCCATTTGCCCCACTTCCACCAAT  
TCCATTTCCAAGATTTCCATGGTTTAGACGTAATTTTCCTATTCCAATACCTGAATCTGCCCC  
TACAACTCCCCTTCCTAGCGAAAAGTAAACAAGAAGGATAAGTCACGATAAACCTGGTCACCT  
GAAATTGAAATTGAGCCACTTCCTTGAAGAATCAAAATTCCTGTTAATAAAAGAAAAACAAAT  
GTAATTGAAATAGCACACAGCATTCTCTAGTCAATATCTTTAGTGATCTTCTTTAATAAACAT  
GAAAGCAAAGATTTTGGTTTCTTAATTTCCACA

464/615

**FIGURE 460**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA71290  
><subunit 1 of 1, 85 aa, 1 stop  
><MW: 9700, pI: 9.55, NX(S/T): 0  
MKKVLLLLITAILAVAVGFPVSQDQEREKRSISDSDELASGFFVFPYPYPERPLPPIPFPRFPW  
FRRNFPIPIPIESAPTTPLPSEK

**Important features of the protein:****Signal peptide:**

amino acids 1-17

**Homologous region to B3-hordein:**

amino acids 47-85

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**N-glycosylation sites:**

Amino acids 57-61;210-214;220-224;318-322;428-432;472-476

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 80-84

**N-myristoylation sites:**Amino acids 3-9;20-29;39-48;152-161;161-170;262-271;358-364;  
538-544;560-566;637-643**Zinc carboxypeptidases, zinc-binding region 2 signature:**

Amino acids 498-509

**Zinc carboxypeptidases:**

Amino acids 391-411

465/615

**FIGURE 461**

AGCAGGAGCAGGAGAGGGGACA**ATG**GAAAGCTGCCCCGTCCAGGTTTCATGTTCTCTTATTTCTC  
CTCACGTGTGAGCTGGCTGCAGAAGTTGCTGCAGAAGTTGAGAAATCCTCAGATGGTCCTGGT  
GCTGCCCAGGAACCCACGTGGCTCACAGATGTCCAGCTGCCATGGAATTCATTGCTGCCACT  
GAGGTGGCTGTCATAGGCTTCTTCCAGGATTTAGAAATACCAGCAGTGCCCATACTCCATAGC  
ATGGTGCAAAAATTCCCAGGCGTGTCATTTGGGATCAGCACTGATTCTGAGGTTCTGACACAC  
TACAACATCACTGGGAACACCATCTGCCTCTTTCGCCTGGTAGACAATGAACAACCTGAATTTA  
GAGGACGAAGACATTGAAAGCATTGATGCCACCAAATTGAGCCGTTTCATTGAGATCAACAGC  
CTCCACATGGTGACAGAGTACAACCCCTGTGACTGTGATTGGGTTATTCAACAGCGTAATTCAG  
ATTCATCTCCTCCTGATAATGAACAAGGCCTCCCCAGAGTATGAAGAGAACATGCACAGATAC  
CAGAAGGCAGCCAAGCTCTTCCAGGGGAAGATTCTCTTTATTCTGGTGGACAGTGGTATGAAA  
GAAAATGGGAAGGTGATATCATTTTTCAAACCTAAAGGAGTCTCAACTGCCAGCTTTGGCAATT  
TACCAGACTCTAGATGACGAGTGGGATACACTGCCCACAGCAGAAAGTTTCCGTAGAGCATGTG  
CAAACTTTTGTGATGGATTCCCTAAGTGGAATTTGTTGAAAGAAAATCGTGAATCAGAAGGA  
AAGACTCCAAAGGTGGAACCT**TGAC**TTCTCCTTGGAACCTACATATGGCCAAGTATCTACTTTA  
TGCAAAGTAAAAAGGCACAACCTCAAATCTCAGAGACACTAAACAACAGGATCACTAGGCCTGC  
CAACCACACACACACGCACGTGCACACACGCACGCACGCGTGCACACACACACGCGCACACAC  
ACACACACACAGAGCTTCATTTCTGTCTTAAATCTCGTTTTCTCTTCTTCTTCTTTTAA  
TTTCATATCCTCACTCCCTATCCAATTTCTTCTTATCGTGCATTCATACTCTGTAAGCCCAT  
CTGTAACACACCTAGATCAAGGCTTTAAGAGACTCACTGTGATGCCTCTATGAAAGAGAGGCA  
TTCCTAGAGAAAGATTGTTCCAATTTGTCATTTAATATCAAGTTTGTATACTGCACATGACTT  
ACACACAACATAGTTCCTGCTCTTTAAGGTTACCTAAGGGTTGAAACTCTACCTTCTTTCAT  
AAGCACATGTCCGTCTCTGACTCAGGATCAAAAACCAAAGGATGGTTTTAAACACCTTTGTGA  
AATTGTCTTTTTGCCAGAAGTTAAAGGCTGTCTCCAAGTCCCTGAACTCAGCAGAAATAGACC  
ATGTGAAAACCTCCATGCTTGGTTAGCATCTCCAACCTCCCTATGTAAATCAACAACCTGCATAA  
TAAATAAAAGGCAATCATGTTATA

466/615

**FIGURE 462**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76401
><subunit 1 of 1, 273 aa, 1 stop
><MW: 30480, pI: 4.60, NX(S/T): 1
MEAAPSREFLLFLLTCELAEEVAAEVEKSSDGPAAQEPTWLTDVPAAMEFIAATEVAVIGF
FQDLEIPAVPILHSMVQKFPGVSEFGISTDSEVLTHYNITGNTICLFRLLVDNEQLNLEDEDIES
IDATKLSRFIEINSLHMTYENPVTVIGLFNSVIQIHLLIMNKASPEYEENMHRYQKAAKLF
QGKILFILVDSGMKENGKVISFFKLKESQLPALAIYQTLDDDEWDTLPTAEVSVEHVQNFCDGF
LSGKLLKENRESEGKTPKVEL
```

**Signal peptide:**

amino acids 1-20

**Transmembrane domain:**

amino acids 143-162

467/615

**FIGURE 463**

CTCGCTTCTTCCTTCTGGATGGGGGGCCAGGGGGCCAGGAGAGTATAAAGGCGATGTGGAGG  
GTGCCCCGGCACAACCAGACGCCCAGTCACAGGCGAGAGCCCTGGG**ATG**CACCGGCCAGAGGCC  
ATGCTGCTGCTGCTCACGCTTGCCCTCCTGGGGGGCCCCACCTGGGCAGGGAAGATGTATGGC  
CCTGGAGGAGGCAAGTATTTTCAGCACCACTGAAGACTACGACCATGAAATCACAGGGCTGCGG  
GTGTCTGTAGGTCTTCTCCTGGTGAAAAGTGTCCAGGTGAACTTGGAGACTCCTGGGACGTG  
AAACTGGGAGCCTTAGGTGGGAATACCCAGGAAGTCACCCTGCAGCCAGGCGAATACATCACA  
AAAGTCTTTGTGCGCTTCCAAGCTTTCCTCCGGGGTATGGTCATGTACACCAGCAAGGACCGC  
TATTTCTATTTTGGGAAGCTTGATGGCCAGATCTCCTCTGCCTACCCAGCCAAGAGGGGCAG  
GTGCTGGTGGGCATCTATGGCCAGTATCAACTCCTTGGCATCAAGAGCATTGGCTTTGAATGG  
AATTATCCACTAGAGGAGCCGACCACTGAGCCACCAGTTAATCTCACATACTCAGCAAACCTCA  
CCCGTGGGTGCGT**TAG**GGTGGGGTATGGGGCCATCCGAGCTGAGGCCATCTGTGTGGTGGTGGC  
TGATGGTACTGGAGTAACTGAGTCGGGACGCTGAATCTGAATCCACCAATAAATAAAGCTTCT  
GCAGAAAA

468/615

**FIGURE 464**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76541
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19600, pI: 5.89, NX(S/T): 1
MHRPEAMLLLLTLALLGGPTWAGKMYGPGGGKYFSTTEDYDHEITGLRVSVGLLLVKSQVVKL
GDSWDVKLGALGGNTQEVTLQPGYITKVFVAFQAFLRGMVMTSKDRYFYFGKLDGQISSAY
PSQEGQVLVGIYGQYQLLGIKSIGFEWNYPLEEPTTEPPVNLTYSANSPVGR
```

**Signal peptide:**  
amino acids 1-22

469/615

**FIGURE 465**

CGGACGCGTGGGTCCGGCGGCCTGAGGCTGCACCGGGCACGGGTGGGCCGAATCCAGCCTGGGCGGAGCCGGAG  
TTGCGAGCCGCTGCCTAGAGGCCGAGGAGCTCACAGCTATGGGCTGGAGCCCCGGAGAGCTCGGGGGACCCCGT  
TGCTGCTGCTGCTACTACTGCTGCTGCTCTGGCCAGTGCCAGGCGCCGGGGTGCTTCAAGGACATATCCCTGGGC  
AGCCAGTCACCCCGCACTGGGTCTGGATGGACAACCCTGGCGCACCGTCAGCCTGGAGGAGCCGGTCTCGAAGC  
CAGACATGGGGCTGGTGGCCCTGGAGGCTGAAGGCCAGGAGCTCCTGCTGAGCTGGAGAAGAACCACAGGCTGC  
TGGCCCCAGGATACATAGAAAACCCACTACGGCCCCAGATGGGCAGCCAGTGGTGCTGGCCCCAACACACGGATC  
ATTGCCACTACCAAGGGCGAGTAAGGGGCTTCCCCGACTCCTGGGTAGTCTCTGCACCTGCTCTGGGATGAGTG  
GCCTGATCACCCCTCAGCAGGAATGCCAGCTATTATCTGCGTCCCTGGCCACCCCGGGGCTCCAAGGACTTCTCAA  
CCCACGAGATCTTTTCGGATGGAGCAGCTGCTCACCTGGAAAGGAACCTGTGGCCACAGGGATCCTGGGAACAAAG  
CGGGCATGACCAGCCTTCTTGGTGGTCCCCAGAGCAGGGGCGAGGCGAGAAGCGCGCAGGACCCGGAAGTACCTGG  
AACTGTACATTGTGGCAGACCACACCCTGTCTTGACTCGGCACCGAACTTGAACCACACCAACAGCGTCTCC  
TGGAAAGTCGCCAACTACGTGGACAGCTTCTCAGGACTCTGGACATTCAAGTGCGCGTGAACGGCCCTGGAGGTGT  
GGACCGAGCGGGACCGCAGCCGCGTCACGCAGGACGCCAACGCCACGCTCTGGGCTTCTGCACTGGCGCCGGG  
GGCTGTGGGCGCAGCGGCCCCACGACTCCGCGCAGCTGCTCACGGGCGCGCCTTCCAGGGCGCCACAGTGGGCC  
TGGCGCCCGTCGAGGGCATGTGCCGCGCCGAGAGCTCGGGAGGCGTGAGCACGGACCACTCGGAGCTCCCATCG  
GCGCCGACGCCACCATGGCCCATGAGATCGGCCACAGCCTCGGCCTCAGCCACGACCCCGACGGCTGCTGCGTGG  
AGGCTGCGGCGGAGTCCGGAGGCTGCGTCATGGCTGCGGCCACCGGGCACCCGTTTCCGCGCGTGTTCAGCGCCT  
GCAGCCGCGCCAGCTGCGCGCCTTCTTCCGCAAGGGGGCGCGCTTGCCCTCTCCAATGCCCGGACCCCGGAC  
TCCCGGTGCCGCGCGCGCTCTGCGGGAACGGCTTCGTGGAAGCGGGCGAGGAGTGTGACTGCGGCCCTGGCCAGG  
AGTGCCGCGACCTCTGCTGCTTGTCTCACAACCTGCTCGCTGCGCCCGGGGGCCAGTGCGCCACGGGGACTGCT  
GCGTGCGCTGCTGCTGAAGCCGGCTGGAGCGCTGTGCCGCCAGGCCATGGGTGACTGTGACCTCCCTGAGTTTT  
GCACGGGCACCTCCTCCCACTGTCCCCAGACGTTTACCTACTGGACGGCTCACCCCTGTGCCAGGGGCGAGTGGCT  
ACTGCTGGGATGGCGCATGTCCACGCTGGAGCAGCAGTGCCAGCAGCTCTGGGGGCTGGCTCCCACCCAGCTC  
CCGAGGCTGTGTTCCAGGTGGTGAACCTCTGCGGAGATGCTCATGGAACTGCGGCCAGGACCGGAGGGCCACT  
TCTGCCCCGTGTGACGGGAGGGATGCCCTGTGTGGGAAGCTGCAGTGCCAGGGTGGAAAGCCAGCCTGTCTGCGAC  
CGCACATGGTGCCAGTGGACTCTACCGTTACCTAGATGGCCAGGAAGTGACTTGTGGGGAGCCTTGGCACTCC  
CCAGTGCCCAGCTGGACCTGCTTGGCCTGGGCTGGTAGAGCCAGGCACCCAGTGTGGACCTAGAAATGGTGTGCC  
AGAGCAGGCGCTGCAGGAAGATGCCTTCCAGGAGCTTCAGCGCTGCCTGACTGCCTGCCACAGCCACGGGGTTT  
GCAATAGCAACCATAACTGCCACTGTGCTCCAGGCTGGGCTCCACCCTTCTGTGACAAGCCAGGCTTTGGTGGCA  
GCATGGACAGTGGCCCTGTGACGGCTGAAAACCATGACACCTTCTGCTGGCCATGCTCCTCAGCGTCTGCTGCTGC  
CTCTGCTCCCAGGGGCGCGCCTGGCCTGGTGTGTTGCTACCGACTCCAGGAGCCCATCTGACGCGATGCAGTGGG  
GCTGCAGAAGGGACCCTGCGTGCAGTGGCCCCAAAGATGGCCACACAGGGACACCCCTGGGCGGCGTTCACC  
CCATGGAGTTGGGCCCCACAGCCACTGGACAGCCCTGGCCCTGGACCTGAGAACTCTCATGAGCCAGCAGCC  
ACCCTGAGAAGCCTCTGCCAGCAGTCTGCGCTGACCCCCAAGCAGATCAAGTCCAGATGCCAAGATCCTGCCTCT  
GGTGAAGGTAGCTCCTAAAATGAACAGATTTAAAGACAGGTGGCCACTGACAGCCACTCCAGGAACCTGAACTG  
CAGGGGCAGAGCCAGTGAATCACCGGACCTCCAGCACCTGCAGGCAGCTTGAAGTTTCTTCCCGAGTGGAGCT  
TCGACCCACCCACTCCAGGAACCCAGAGCCACATTAGAAGTTCTGAGGGCTGGAGAACACTGCTTGGGCACACT  
CTCCAGCTCAATAAACCATCAGTCCCAGAAGCAAAGGTACACAGCCCTGACCTCCCTCACCAGTGGAGGCTGG  
GTAGTGTGGCCATCCAAAAGGGCTCTGTCTGGGAGTCTGGTGTGTCTCTACATGCAATTTCCACGGACCCA  
GCTCTGTGGAGGGCATGACTGCTGGCCAGAAGCTAGTGGTCTGGGGCCCTATGGTTCGACTGAGTCCACACTCC  
CCTGCAGCCTGGCTGGCCTCTGCAACAAACATAATTTTGGGGACCTTCTTCTGTTTCTTCCACCCCTGTCTT  
CTCCCCTAGGTGGTTCCTGAGCCCCACCCCCAATCCAGTGCTACACCTGAGGTTCTGGAGCTCAGAATCTGAC  
AGCCTCTCCCCATTCTGTGTGTGTCCGGGGGACAGAGGGAACCATTTAAGAAAAGATACCAAAGTAGAAGTCAA  
AAGAAAGACATGTTGGCTATAGGCGTGGTGGCTCATGCCTATAATCCCAGCACTTTGGGAAGCCGGGGTAGGAGG  
ATCACAGAGGGCCAGCAGGTCCACACCAGCCTGGGCAACACAGCAAGACACCGCATCTACAGAAAAATTTAAAA  
TTAGCTGGGCGTGGTGGTGTACCTGTAGGCCTAGTGTCTCAGGAGGCTGAAGCAGGAGGATCACTTGAGCCTG  
AGTTCAACACTGCAGTGAGCTATGGTGGCACCACTGCACTCCAGCCTGGGTGACAGAGCAAGACCCCTGTCTCTAA  
AATAAATTTTAAAGGACTTAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAGAAAA



470/615

**FIGURE 466**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76788  
><subunit 1 of 1, 813 aa, 1 stop  
><MW: 87739, pI: 6.94, NX(S/T): 5

MGWRPRRARGTFLLLLLLLLLLWFPVPGAGVLQGHIPGQPVTPHWVLDGQPWRTVSLEEPVSKPDMGLVALEAEGQ  
ELLLELEKHNHRLAPGYIETHYGPDGQPVVLAPNHTDHCHYQGRVRFPSWVVLCTCSGMSGSLITLSRNASYLL  
RPWPPRGSKDFSTHEIFRMEQLLTWKGTGHRDPGNKAGMTSLPGGPQSRGRREARRTRKYLELYIVADHTLFLT  
RHRNLNHTKQRLLEVANYVDQLRLTDIQVALTGLEVWTERDRSRVTQDANATLWAFLOWRRGLWAQRPHDSAQL  
LTGRAFGGATVGLAPVEGMCRAESSGGVSTDHSELPIGAAATMAHEIGHSLGLSHDPDGCCVEAAESGGCVMAA  
ATGHPFPRVFSACSRRLRAFFRKGGGACLSNAPDPGLPVPPALCGNGFVEAGEECDGPGQECRDLCFAHNCS  
LRPGAQCAHGDCCVRCLLKPAGALCRQAMGDCDLPEFCTGTSSHCPPDVYLLDGSPCARGSGYCWDGACPTLEQQ  
CQQLWGPGSHPAPEACFQVVSAGDAHGNCQDSEGHFLPCAGRDALCGKLQCQGGKPSLLAPHMVPVDSTVHLD  
GQEVTCRGALALPSAQLDLLGLGLVEPGTQCGPRMVCQSRRCRKNAFQELQRCLTACHSHGVCNSNHNCHCAPGW  
APPFCDKPGFGGSMDSGPVQAEHDTFLAMLLSVLLPLLPGAGLAWCCYRLPGAHLQRCSWGCRDRPACSGPKD  
GPHRDHPLGGVHPMELGPTATGQPWPLDPENSHPSSHPEKPLPAVSPDPQADQVQMPRSLIW

**Important features of the protein:****Signal peptide:**

Amino acids 1-27

**Transmembrane domain:**

Amino acids 702-720

**N-glycosylation sites:**

Amino acids 109-113;145-149;231-235;276-280;448-452

**Tyrosine kinase phosphorylation site:**

Amino acids 236-244

**N-myristoylation sites:**

Amino acids 29-35;185-191;195-201;308-314;318-324;326-332;338-344;370-376;  
400-406;402-408;454-460;504-510;510-516;517-523;580-586;  
601-607;661-667;687-693;717-723;719-725

**Amidation site:**

Amino acids 200-204

**Neutral zinc metallopeptidases, zinc-binding region signature:**

Amino acids 342-352

471/615

**FIGURE 467**

[illegible]

472/615

**FIGURE 468**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA77623  
><subunit 1 of 1, 97 aa, 1 stop  
><MW: 10160, pI: 6.56, NX(S/T): 0  
MQLGTGLLLAAVLSLQLAAAEAIWCHQCTGFGGCSHGSRCLRDSTHCVTTATRVLSNTEDLPL  
VTKMCHIGCPDIPSLGLGPYYSIACCQTSLCNHD

**Important features of the protein:****Signal peptide:**

amino acids 1-20

**N-myristoylation sites.**

amino acids 6-11 and 33-38

**Prokaryotic membrane lipoprotein lipid attachment sites.**

amino acids 24-34 and 78-88

473/615

**FIGURE 469**

**CATG**GAGCCTCTTGCAGCTTACCCGCTAAAATGTTCCGGGCCCAGAGCAAAGGTATTTGCAGT  
TTTGCTGTCTATAGTTCTATGCACAGTAACGCTATTTCTTCTACAACATAAAATTCCTCAAACC  
TAAAATCAACAGCTTTTATGCCTTTGAAGTGAAGGATGCAAAGGAAGAAGTGTCTCTGGA  
AAAGTATAAAGGCAAAGTTTCACTAGTTGTAAACGTGGCCAGTGACTGCCAACTCACAGACAG  
AAATTACTTAGGGCTGAAGGAAGTGCACAAAGAGTTTGGACCATCCCACTTCAGCGTGTGGC  
TTTTCCCTGCAATCAGTTTGGAGAATCGGAGCCCCGCCAAGCAAGGAAGTAGAATCTTTTGC  
AAGAAAAAAGTACGGAGTAACTTTCCCATCTTCCACAAGATTAAGATTCTAGGATCTGAAGG  
AGAACCTGCATTTAGATTTCTTGTGATTCTTCAAAGAAGGAACCAAGGTGGAATTTTGGAA  
GTATCTTGTCAACCCTGAGGGTCAAGTTGTGAAGTTCTGGAGGCCAGAGGAGCCCATTGAAGT  
CATCAGGCCTGACATAGCAGCTCTGGTTAGACAAGTGATCATAAAAAAGAAAGAGGATCT**ATG**  
**A**GAATGCCATTGCGTTTCTAATAGAACAGAGAAATGTCTCCATGAGGGTTGGTCTCATTTTA  
AACATTTTTTTTTTGGAGACAGTGTCTCACTCTGTCACCCAGGCTGGAGTGCAGTAGTGCCTT  
CTCAGCTCATTTGCAACCTCTGCCTTTTTTAAACATGCTATTAAATGTGGCAATGAAGGATTTTT  
TTTTAATGTTATCTTGCTATTAAGTGGTAATGAATGTTCCCAGGATGAGGATGTTACCCAAAG  
CAAAAATCAAGAGTAGCCAAAGAATCAACATGAAATATATTAAGTACTTCCCTCTGACCATACT  
AAAGAATTCAGAATACACAGTGACCAATGTGCCTCAATATCTTATTGTTCAACTTGACATTTT  
CTAGGACTGTACTTGATGAAAATGCCAACACACTAGACCACTCTTTGGATTCAAGAGCACTGT  
GTATGACTGAAATTTCTGGAATAACTGTAAATGGTTATGTTAATGGAATAAAACACAAATGTT  
GAAAAATGTAAATATATATACATAGATTCAAATCCTTATATATGTATGCTTGTTTTGTGTAC  
AGGATTTTGTTTTTTCTTTTTAAGTACAGGTTCCCTAGTGTTTTACTATAACTGTCACTATGTA  
TGTAAGTACATATATAAATAGTCATTTATAAATGACCGTATTATAACATTTGAAAAAGTCTT  
CATCAAAAAAAAAAAAAA

474/615

**FIGURE 470**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA80136
><subunit 1 of 1, 209 aa, 1 stop
><MW: 23909, pI: 9.68, NX(S/T): 0
MEPLAAYPLKCSGPRAKVFAVLLSIVLCTVTLFLLQLKFLKPKINSFYAFEVKDAKGRTVSLE
KYKGKVSLLVNVASDCQLTDRNYLGLKELHKEFGPSHFSVLAFFPCNQFGESEPRPSKEVESFA
RKNYGVTFPIFHKIKILGSEGEPAFRFLVDSSKKEPRWNFWKYLVNPEGQVVKFWRPEEPIEV
IRPDIAALVRQVIIKKKEDL
```

**Important features of the protein:****Signal peptide:**

amino acids 1-31

**Glutathione peroxidases signature 2.**

amino acids 104-112

**Glutathione peroxidases.**

amino acids 57-82

475/615

**FIGURE 471**

GCCCTAACCTTCCCAGGGCTCAGCTCTTTGGAGCTGCCATTCTCCGGCTGCGAGAAAGGACGCGCGCCCTGCG  
TCGGGCGAAGAAAAGAAGCAAACTTGTGCGGAGGGTTTCGTCATCAACCTCCTTCCCGCAAACCTAAACCTCCT  
GCCGGGGCCATCCCTAGACAGAGGAAAGTTCTGCGAGAGCCGACCAGCCCTAGTGGATCTGGGGCAGGCAGCGGC  
GCTGGCTGTGGAATTAGATCTGTTTTGAACCCAGTGGAGCGCATCGCTGGGGCTCGGAAGTCACCGTCCGCGGGC  
ACCGGGTTGGCGCTGCCCCAGTGAACCGACAGTTTGGCAGCCTCGGCTGCAAGTGGCCTCTCCTCCCCGCGGTT  
GTTGTTCACTGTGCGGGTGAAGGCTGCGAGTGTGGCAAGTTGCAAAAGAGAGCCTCAGAGGTCCGAAGAGCGCTGCG  
CTCCTACTCGCGTTGCGTCTCTCTCTCGGTTCCCTACTGTGAAATCGCAGCGACATTTACAAAGGCCTCCG  
GGTCTACCGAGACCGATCCGAGCGTTTGGCCCGGTGCTGCCATATGTCATCGGGAGCCCCGAGCACCGGCGAA  
GGACTGGCGGGTGGGGTAGGGAGGTGGCGGCGGCGGCATGGCGAGGTTCCCGAAGGCCGACCTGGCCGCTGCAGG  
AGTTATGTTACTTTGCCACTTCTTCACGGACCAAGTTTCAGTTTCGCGGATGGGAAACCCGGAGACCAATCCTTGA  
TTGGCAGTATGGAGTTACTCAGGCCTTCCCTCACACAGAGGAGGAGGTGGAAGTTGATTACACGCGTACAGCCA  
CAGGTGGAAAAGAACTTGGACTTCTCAAGGCGGTAGACACGAACCGAGCAAGCGTCGGCCAAGACTCTCCTGA  
GCCAGAAGCTTCACAGACCTGCTGCTGGATGATGGGCAGGACAATAACACTCAGATCGAGGAGGATACAGACCA  
CAATTACTATATATCTCGAATATATGGTCCATCTGATTCTGCCAGCCGGGATTTATGGGTGAACATAGACCAAT  
GGAAAAGATAAAGTGAAGATTCATGGAATATTGTCCAATACTCATCGCAAGCTGCAAGAGTGAATCTGTCCTT  
CGATTTTCCATTTTATGGCCACTTCCCTACGTGAAATCACTGTGGCAACCGGGGTTTCATATACACTGGAGAAGT  
CGTACATCGAATGCTAACAGCCACACAGTACATAGCACCTTTAATGGCAATTTTCGATCCCAGTGTATCCAGAAA  
TTCAACTGTGAGATATTTGATAATGGCACAGCACTTGTGGTCCAGTGGGACCATGTACATCTCCAGGATAATTA  
TAACCTGGGAAGCTTCACATTCCAGGCAACCTGCTCATGGATGGACGAATCATCTTTGGATACAAAGAAATTC  
TGTCTTGGTCACACAGATAAGTTCAACCAATCATCCAGTGAAGTTCGGACTGTCCGATGCATTTGTGCTTGTCCA  
CAGGATCCAACAAATTTCCCAATGTTGGAAGAAGAACAATTTATGAATACCACCGAGTAGAGCTACAAATGTCAAA  
AATTACCAACATTTCCGGCTGTGGAGATGACCCCATACCCACATGCCTCCAGTTTAAACAGATGTGGCCCCCTGTGT  
ATCTTCTCAGATTGGCTTCAACTGCAGTTGGTGTAGTAACTTCAAAGATGTTCCAGTGGATTTGATCGTCATCG  
GCAGGACTGGGTGGACAGTGGATGCCCTGAAGAGTCAAAAGAGAAGATGTGTGAGAATACAGAACCAGTGGAAAC  
TTCTTCTCGAACCACCACAACCGTAGGAGCGACAACCACCCAGTTTCAGGGTCCTAACTACCACCAGAAGAGCAGT  
GACTTCTCAGTTTCCCACAGCCTCCCTACAGAAGATGATACCAAGATAGCACTACATCTAAAAGATAATGGAGC  
TTCTACAGATGACAGTGCAGCTGAGAAGAAAGGGGAACCCCTCCACGCTGGCCTCATCATTGGAATCCTCATCCT  
GGTCTCTATTGTAGCCACAGCCATTCTTGTGACAGTCTATATGTATCACCACCCAACATCAGCAGCCAGCATCTT  
CTTTATTGAGAGACGCCCCAAGCAGATGGCCTGCGATGAAGTTTAGAAGAGGCTCTGGACATCCTGCCTATGCTGA  
AGTTGAACCAAGTTGGAGAGAAAGAAGGCTTTATTGTATCAGAGCAGTGCTTAAATTTCTAGGACAGAACAACACC  
AGTACTGGTTTACAGGTGTTAAGACTAAAATTTTGCCTATACCTTTAAGACAAACAAACACACACACAAAC  
AAGCTCTAAGCTGCTGTAGCCTGAAGAAGACAAGATTTCTGGACAAGCTCAGCCCAGGAAACAAAGGGTAAACAA  
AAAACATAAACTTATACAAGATACCATTTACACTGAACATAGAATTCCTAGTGGAAATGTCATCTATAGTTCACT  
CGGAACATCTCCCGTGGACTTATCTGAAGTATGACAAGATTATAATGCTTTTGGCTTAGGTGCAGGTTGCAAG  
GGATCAGAAAAAAAATCATAATAAAGCTTTAGTTTCATGAGGG

476/615

**FIGURE 472**

MARFPKADLAAAGVMLLCHFFTDQFQFADGKPGDQILDWQYGVTAQAFPHTEEEVEVD SHAYSH  
RWKRNLDFLKAVDTNRASVGQDSPEPRSFTDLLLDDGQDNNTQIEEDTDHNYIISRIYGPSDS  
ASRDLWVNIDQMEKDKVKI HGILSNTHRQAARVNLSFDFPFYGHFLREITVATGGFIYTGEVV  
HRMLTATQYIAPLMANFDPSVSRNSTVRYFDNGTALVVQWDHVHLQDNYNLGSFTFQATLLMD  
GRIIFGYKEIPVLVTQISSTNHPVKVGLS DAFVVVHRIQQIPNVRRTIY EYHRVELQMSKIT  
NISAVEMTPLPTCLQFNRCGPCVSSQIGFNC SWCSKLQRCSSGFDRHRQDWVDSGCPEESKEK  
MCENTE PVETSSRTTTTVGATTTQFRVLT TTTT RRAVTSQFPTSLPTEDDTKIALHLKDNGASTD  
DSAAEKKGGTLHAGLIIGILILVLIVATAILVTVMYHHPTSAASIFFIERRPSRWPAMKFRR  
GSGHPAYAEVEPVGEKEGFIVSEQC

**Important features of the protein:****Transmembrane domain:**

amino acids 454-478

**N-glycosylation sites.**

amino acids 103-107, 160-164, 213-217, 221-225, 316-320, 345-349

**cAMP- and cGMP-dependent protein kinase phosphorylation sites.**

amino acids 297-301, 492-496, 503-507

**N-myristoylation sites.**amino acids 42-48, 100-106, 147-153, 279-285, 397-403, 450-456,  
455-461

477/615

**FIGURE 473**

CGCGGAGCCCTGCGCTGGGAGGTGCACGGTGTGCACGCTGGACTGGACCCCCATGCAACCCCG  
CGCCCTGCGCCTTAACCAGGACTGCTCCGCGCGCCCCTGAGCCTCGGGCTCCGGCCCGGACCT  
GCAGCCTCCCAGGTGGCTGGGAAGAACTCTCCAACAATAAATACATTTGATAAGAAAG**ATGGC**  
TTTAAAAGTGCTACTAGAACAAAGAGAAAACGTTTTTCACTCTTTTAGTATTACTAGGCTATTT  
GTCATGTAAAGTGACTTGTGAATCAGGAGACTGTAGACAGCAAGAATTCAGGGATCGGTCTGG  
AACTGTGTTCCCTGCAACCAGTGTGGGCCAGGCATGGAGTTGTCTAAGGAATGTGGCTTCGGC  
TATGGGGAGGATGCACAGTGTGTGACGTGCCGGCTGCACAGGTTCAAGGAGGACTGGGGCTTC  
CAGAAATGCAAGCCCTGTCTGGACTGCGCAGTGGTGAACCGCTTTCAGAAGGCAAATTTGTTCA  
GCCACCAGTGATGCCATCTGCGGGGACTGCTTGCCAGGATTTTATAGGAAGACGAACTTGTC  
GGCTTTCAAGACATGGAGTGTGTGCCTTGTGGAGACCCTCCTCCTCCTTACGAACCGCACTGT  
GCCAGCAAGGTCAACCTCGTGAAGATCGCGTCCACGGCCTCCAGCCCACGGGACACGGCGCTG  
GCTGCCGTTATCTGCAGCGCTCTGGCCACCGTCTGCTGGCCCTGCTCATCCTCTGTGTCTC  
TATTGTAAGAGACAGTTTATGGAGAAGAAACCCAGCTGGTCTCTGCGGTCGCAGGACATTCAG  
TACAACGGCTCTGAGCTGTCTGTTTTGACAGACCTCAGCTCCACGAATATGCCACAGAGCC  
TGCTGCCAGTGCCGCCGTGACTCAGTGCAGACCTGCGGGCCGGTGCGCTTGCTCCCATCCATG  
TGCTGTGAGGAGGCCTGCAGCCCCAACCCGGCGACTCTTGGTTGTGGGGTGCATTCTGCAGCC  
AGTCTTCAGGCAAGAAACGCAGGCCAGCCGGGAGATGGTGCCGACTTTCTTCGGATCCCTC  
ACGCAGTCCATCTGTGGCGAGTTTTTCAGATGCCTGGCCTCTGATGCAGAATCCCATGGGTGGT  
GACAACATCTCTTTTTGTGACTCTTATCCTGAACTCACTGGAGAAGACATTCATTCTCTCAAT  
CCAGAACTTGAAAGCTCAACGTCTTTGGATTCAAATAGCAGTCAAGATTTGGTTGGTGGGGCT  
GTTCCAGTCCAGTCTCATTCTGAAAACTTTACAGCAGCTACTGATTTATCTAGATATAACAAC  
ACACTGGTAGAATCAGCATCAACTCAGGATGCACTAACTATGAGAAGCCAGCTAGATCAGGAG  
AGTGGCGCTGTCTATCCACCCAGCCACTCAGACGTCCCTCCAGGAAGCT**TAA**AGAACCTGCTTC  
TTTCTGCAGTAGAAGCGTGTGCTGGAACCCAAAGAGTACTCCTTTGTTAGGCTTATGGACTGA  
GCAGTCTGGACCTTGCATGGCTTCTGGGGCAAATAAATCTGAACCAAACCTGACGGCATTG  
AAGCCTTTCAGCCAGTTGCTTCTGAGCCAGACCAGCTGTAAGCTGAAACCTCAATGAATAACA  
AGAAAAGACTCCAGGCCGACTCATGATACTCTGCATCTTTCCTACATGAGAAGCTTCTCTGCCAC  
AAAAGTGACTTCAAAGACTGATGGGTTGAGCTGGCAGCCTATGAGATTGTGGACATATAACAA  
GAAACAGAAATGCCCTCATGCTTATTTTCATGGTGATTGTGGTTTACAAGACTGAAGACCCA  
GAGTATACTTTTTCTTTCCAGAAATAATTTACATACCGCCTATGAAATATCAGATAAATTACCT  
TAGCTTTTATGTAGAATGGGTTCAAAGTGAGTGTTTCTATTTGAGAAGGACACTTTTTCATC  
ATCTAAACTGATTTCGCATAGGTGGTTAGAATGGCCCTCATATTGCCTGCCTAAATCTTGGGTT  
TATTAGATGAAGTTTACTGAATCAGAGGAATCAGACAGAGGAGGATAGCTCTTTCAGAAATCC  
ACACTTCTGACCTCAGCCTCGGTCTCATGAACACCCGCTGATCTCAGGAGAACACCTGGGCTA  
GGAAATGTGGTCGAGAAAGGGCAGCCCATTGCCAGAAATTAACACATATTGTAGAGACTTGTA  
TGCAAAGGTTGGCATATTTATATGAAAATTAGTTGCTATAGAAACATTTGTTGCATCTGTCCC  
TCTGCCTGAGCTTAGAAGGTTATAGAAAAAGGGTATTTATAAACATAAATGACCTTTTACTTG  
CATTGTATCTTATACTAAAGGCTTTAGAAATTACAACATATCAGGTTCCCTACTACTGAAGT  
AGCCTTCCGTGAGAACACACCACATGTTAGGACTAGAAGAAAATGCACAATTTGTAGGGGTTT  
GGATGAAGCAGCTGTAACCTGCCCTAGTGTAGTTTGACCAGGACATTGTCGTGCTCCTCCAAT  
TGTGTAAGATTAGTTAGCACATCATCTCCTACTTTAGCCATCCGGTGTTGGATTTAAGAGGAC  
GGTGCTTCTTTCTATTAAAGTGCTCCATCCCCTACCATCTACACATTAGCATTGTCTCTAGAG  
CTAAGACAGAAATTAACCCCGTTCAGTCACAAAGCAGGGAATGGTTCATTTACTCTTAATCTT  
TATGCCCTGGAGAAGACCTACTTGAACAGGGCATATTTTTTAGACTTCTGAACATCAGTATGT  
TCGAGGGTACTATGATATTTTGGTTTGGAAATTGCCCTGCCCAAGTCACTGTCTTTAACTTTT  
AACTGAATATTAAATGTATCTGTCTTTCCT



478/615

**FIGURE 474**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA84210
><subunit 1 of 1, 417 aa, 1 stop
><MW: 45305, pI: 5.12, NX(S/T): 6
MALKVLEQEKTFFTLVLLGYLSCKVTCESGDCRQQEFRDRSGNCVPCNQCGPGMELSK
ECGFGYGEDAQCVTCRLHRFKEDWGFQKCKPCLDCAVVNRQKANCSDAICGDCLPG
FYRKTKLVGFQDMECVPCGDP PPPYEPHCASKVNLVKIASTASSPRDTALAAVICSLAT
VLLALLILCVIYCKRQFMEEKPSWSLRSQDIQYNGSELSCFDRPQLHEYAHRACCQCRD
SVQTCGPVRLLPSCCEEACSPNPATLGCQGVHSAASLQARNAGPAGEMVPTFFGSLTQSI
CGEFSDAWPLMQNPMGGDNISFCDSYPELTGEDIHSLNPELESSTSLDSNSSQDLVGGAV
PVQSHSENFATAATDLTRYNNLTVESASTQDALTMRSQLDQESGAVIHPATQTSLQEA
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-25

**Transmembrane domain:**

Amino acids 169-192

**N-glycosylation sites:**

Amino acids 105-109;214-218;319-323;350-354;368-372;379-383

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 200-204;238-242

**Tyrosine kinase phosphorylation site:**

Amino acids 207-214

**N-myristoylation sites:**

Amino acids 55-61;215-221;270-276

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 259-270

**TNFR/NGFR family cysteine-rich region proteins:**

Amino acids 89-96

479/615

**FIGURE 475**

AGCCAGGCAGCACATCACAGCGGGAGGAGCTGTCCCAGGTGGCCCAGCTCAGCAATGGCAATG  
GGGGTCCCCAGAGTCATTCTGCTCTGCCTCTTTGGGGCTGCGCTCTGCCTGACAGGGTCCCAA  
GCCCTGCAGTGCTACAGCTTTGAGCACACCTACTTTGGCCCCTTTGACCTCAGGGCCATGAAG  
CTGCCCAGCATCTCCTGTCTCATGAGTGCTTTGAGGCTATCCTGTCTCTGGACACCGGGTAT  
CGCGCGCCGGTGACCCTGGTGCGGAAGGGCTGCTGGACCGGGCCTCCTGCGGGCCAGACGCAA  
TCGAACCCGGACGCGCTGCCGCCAGACTACTCGGTGGTGCGCGGCTGCACAACCTGACAAATGC  
AACGCCCACCTCATGACTCATGACGCCCTCCCCAACCTGAGCCAAGCACCCGACCCGCCGACG  
CTCAGCGGCGCCGAGTGCTACGCCTGTATCGGGGTCCACCAGGATGACTGCGCTATCGGCAGG  
TCCCGACGAGTCCAGTGTACCAGGACCAGACCGCCTGCTTCCAGGGCAGTGGCAGAATGACA  
GTTGGCAATTTCTCAGTCCCTGTGTACATCAGAACCTGCCACCGGCCCTCCTGCACCACCGAG  
GGCACCACCAGCCCCCTGGACAGCCATCGACCTCCAGGGCTCCTGCTGTGAGGGGTACCTCTGC  
AACAGGAAATCCATGACCCAGCCCTTCACCAGTGCTTCAGCCACCACCCCTCCCCGAGCACTA  
CAGGTCCTGGCCCTGCTCCTCCAGTCCTCCTGCTGGTGGGGCTCTCAGCATAGACCGCCCT  
CCAGGATGCTGGGGACAGGGCTCACACACCTCATTCTTGCTGCTTCAGCCCCTATCACATAGC  
TCACTGGAAAATGATGTTAAAGTAAGAATTGCAAAA

480/615

**FIGURE 476**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA86576

&gt;&lt;subunit 1 of 1, 251 aa, 1 stop

&gt;&lt;MW: 26935, pI: 7.42, NX(S/T): 2

MAMGVPRVILLCLFGAALCLTGSQALQCYSFEHTYFGPFDLRAMKLPSISCPHECFEAILSLD  
TGYRAPVTLVRKGCWTGPPAGQTQSNPDALPPDYSVVRGCTTDKCNAHLMTHDALPNLSQAPD  
PPTLSGAECYACIGVHQDDCAIGRSRRVQCHQDQTACFQSGRMTVGNFSVPVYIIRTCHRPSC  
TTEGTTSPWTAIDLQGSCEGYLCNRKSMTQPFTSASATTPPRALQVLALLLPVLLLVGLSA

**Important features of the protein:****Signal peptide:**

amino acids 1-19

**Transmembrane domain:**

amino acids 233-251

**N-glycosylation sites.**

amino acids 120-124, 174-178

**N-myristoylation sites.**

amino acids 15-21, 84-90

481/615

**FIGURE 477**

CCCACGCGTCCGGGACAGATGAACTTAAAAGAGAAGCTTTAGCTGCCAAAGATTGGGAAAGGGAAAGGACAAAA  
AGACCCCTGGGCTACACGGCGTAGGTGCAGGGTTTCTACTGCTGTTCTTTTATGCTGGGAGCTGTGGCTGTAAC  
CAACTAGGAAATAACGTATGCAGCAGCTATGGCTGTCAGAGAGTTGTGCTTCCCAAGACAAAGGCAAGTCCTGTT  
TCTTTTTCTTTTTGGGGAGTGTCTTGGCAGGTTCTGGGTTTGGACGTTATTTCGGTGACTGAGGAAACAGAGAA  
AGGATCCTTTGTGGTCAATCTGGCAAAGGATCTGGGACTAGCAGAGGGGGAGCTGGCTGCAAGGGGAACCAGGGT  
GGTTTCCGATGATAACAAACAATACCTGCTCCTGGATTACATACCGGGAATTTGCTCACAAATGAGAACTGGA  
CCGAGAGAAGCTGTGTGGCCCTAAAGAGCCCTGTATGCTGTATTTCCAAATTTAATGGATGATCCCTTTCAGAT  
TTACCGGGCTGAGCTGAGAGTCAGGGATATAAATGATCACGCGCCAGTATTTTCAGGACAAAGAAACAGTCTTAAA  
AATATCAGAAAATACAGCTGAAGGGACAGCATTTAGACTAGAAAGAGCACAGGATCCAGATGGAGGACTTAACGG  
TATCCAAAACCTACACGATCAGCCCCAATCTTTTTTCCATATTAACATTAGTGGCGGTGATGAAGGCATGATATA  
TCCAGAGCTAGTGTGGACAAAGCACTGGATCGGGAGGAGCAGGGAGAGCTCAGCTTAACCCCTCACAGCGCTGGA  
TGCTGGGTCTCCATCCAGGTCTGGGACCTCTACTGTACGCATCGTTGTCTTGACGCTCAATGACAATGCCCCACA  
GTTTGGCCAGGCTCTGTATGAGACCCAGGCTCCAGAAAAACAGCCCCATTGGGTTTCTTATTGTTAAGGTATGGGC  
AGAAGATGTAGACTCTGGAGTCAACGCGGAAGTATCCTATTCATTTTTGTATGCTCAGAAAATATTTCGAACGAC  
CTTTCAAATCAATCCTTTTTCTGGGAAATCTTTCTCAGAGAATTGCTTGATTATGAGTTAGTAAATCTTACAA  
AATAAATATACAGGCAATGGACGGTGGAGGCCCTTTCTGCAAGATGTAGGGTTTATGTTGAAGTATTGGACACCAA  
TGACAATCCCCCTGAATGATCGTATCATCATTTTTCCAACCTCTGTTGCTGAGAATTCTCCTGAGACGCCGCTGGC  
TGTTTTTAAGATTAATGACAGAGACTCTGGAGAAAATGGAAAGATGGTTTGCTACATTCAAGAGAATCTGCCATT  
CCTACTAAAACCTTCTGTGGAGAATTTTTACATCCTAATTACAGAAGGCGCGCTGGACAGAGAGATCAGAGCCGA  
GTACAACATCACTATCACCGTCACTGACTTGGGGACACCCAGGCTGAAAACCGAGCACAAACATAACGGTCCCTGGT  
CTCCGACGTCAATGACAACGCCCCCGCTTCAACCAACCTCCTACACCCTGTTGCTCCGCGAGAACAACAGCCC  
CGCCCTGCACATCGGCAGCGTCAGCGCCACAGACAGAGACTCGGGCACCAACGCCAGGTCACCTACTCGCTGCT  
GCCGCCCAAGACCCGCACCTGCCCCCTCGCCTCCCTGGTCTCCATCAACGCGGACAACGGCCACCTGTTCCGCCCT  
CAGGTCGCTGGACTACGAGGCCCTGCAGGCTTTTCAGTTCCGCGTGGGCGCCACAGACCGCGGCTCCCCCGCGCT  
GAGCAGAGAGGCGCTGGTGCGCTGCTGGTGGACGCCAACGACAACCTCGCCCTTCGTGCTGTACCCGCTGCA  
GAACGGCTCCGCGCCCTGCACCGAGCTGGTGCCCCGGGCGGCGAGCCGGGCTACCTGGTGACCAAGGTGGTGGC  
GGTGGACGGCGACTCGGGCCAGAACGCCTGGCTGTGCTACCAGCTGCTCAAGGCCACGGAGCCCGGGCTGTTCCG  
TGTGTGGGCGCACAAATGGGGAGGTGCGCACCGCCAGGCTGCTGAGCGAGCGCGACGCAGCCAAGCACAGGCTCGT  
GGTGCTTGTCAAGGACAATGGCGAGCCTCCTCGCTCGGCCACCGCCACGCTGCACTTGCTCCTGGTGGACGGCTT  
CTCCCAGCCCTACCTGCCTCTCCCGAGGCGGCCCCGGCCAGGCCAGGCCAGGCCAGGCCGACTTGCTCACCGTCTA  
CCTGGTGGTGGCGTTGGCCTCGGTGCTTTCGCTCTTCTCCTCCTCCTCGGTGCTCCTGTTCTGCTGGCGGTGCGGCTGTG  
CAGGAGGAGCAGGGCGGCTCGGTGGGTGCTGCTCGGTGCCCGAGGGTCTTTTCCAGGGCATCTGGTGGACGT  
GAGGGGCGCTGAGACCCTGTCCAGAGCTACAGTATGAGGTGTGCTGACGGGAGGCCCCGGGACCAGTGAGTT  
CAAGTTCTTGAAACAGTTATTTTCGGATATTCAGGCACAGGGCCCTGGGAGGAAGGGTGAAGAAAATTCACCTT  
CCGAAATAGCTTTGGATTTAATATTAGTAAAGTCTGTTTTTAGTTTCATATACTTTTGGTGTGTTACATAGCCA  
TGTTTTCTATTAGTTTACTTTTAAATCTCAAATTTAAGTTATTATGCAACTTCAAGCATTATTTTCAAGTAGTATA  
CCCCGTGGTTTTTACAATGTTTCATCATTTTTTTCATTAAATAACAACCTGGGTTTAATTTAATGAGTATTTTTTT  
CTAAATGATAGTGTTAAGGTTTTAATTCTTTCCAACCTGCCAAGGAATTAATTACTATTATATCTCATTACAGAA  
ATCTGAGTTTTTGATTCATTTTCAGAGCTTGCATCTCATGATTCTAATCACTTCTGTCTATAGTGTACTTGCTCTA  
TTTAAGAAGGCATATCTACATTTCCAAACTCATTCTAACATTCTATATATTTCGTGTTTGAACCATGTCATTTA  
TTTCTACATCATGTATTTAAAAAGAAATATTTCTCTACTACTATGCTCATGACAAAATGAAACAAAGCATATTGT  
GAGCAATACTGAACATCAATAATACCCTTAGTTTATATACTTATTATTTTATCTTTAAGCATGCTACTTTTACTT  
GGCCAATATTTTCTTATGTTAACTTTTGTGATGTATAAACAGACTATGCCTTATAATTGAAATAAAATATATAA  
TCTGCCTGAAAATGAATAAAAATAAAACATTTTGAAATGTGAAAAAAAAAAAAAAAAAAAAA

482/615

**FIGURE 478**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA87976

&gt;&lt;subunit 1 of 1, 800 aa, 1 stop

&gt;&lt;MW: 87621, pI: 4.77, NX(S/T): 7

MAVRELCFPRQRQVLFLEFLFWGVSLAGSGFGGRYSVTEETKGSFVVNLAKDLGLAEGELAARG  
 TRVVSDDNKQYLLLDSTGNLLTNEKLDREKLCGPKEPCMLYFQILMDDPFQIYRAELRVRDI  
 NDHAPVFQDKETVLKISENTAEGTAFRLERAQDPDGLNGIQNYTISPNSFFHINISGGDEGM  
 IYPELVLDKALDREEQGEISLTALTALDGGSPSRSGTSTVRIVVLDVNDNAPQFAQALYETQAP  
 ENSPIGFLIVKVAEDVDSGVNAEVSYSFFDASENIRTTFQINPFSGEIFLRELLDYELVNSY  
 KINIQAAMDGGGLSARCRVLVEVLDTNDNPPELIVSSFSNSVAENSPETPLAVFKINDRDSGEN  
 GKMVCYIQENLPFLKPSVENFYILITEGALDREIRAENITITVTDLGTPLKTEHNITVLV  
 SDVNDNAPAFQTQTSYTLFVRENNSPALHIGSVSATDRDSGTNAQVTYSLPPQDPHLPLASLV  
 SINADNGHLFALRSLDYELQAFEFVRGATDRGSPALSREALVRVLVLDANDNSPFVLYPLQN  
 GSAPCTELVPRAAEPGYLVTKVAVDGDGQNAWLSYQLLKATEPGLFGVWAHNGEVRTARLL  
 SERDAAKHRLVVLVKDNGEPPRSATATLHLLLVDFGFSQPYLPLPEAAPAQAEADLLTVYLV  
 VALASVSSLFLLSVLLFVAVRLCRRSRAASVGRCSVPEGPFPGHLVDVVRGAETLSQSYQYEV  
 LTGGPGTSEFKFLKPVISDIQAQGPGRKGEENSTFRNSFGFNIQ

**Important features of the protein:****Signal peptide:**

amino acids 1-26

**Transmembrane domain:**

amino acids 687-711

**N-glycosylation sites.**

amino acids 169-173, 181-185, 418-422, 436-440, 567-571, 788-792

**Glycosaminoglycan attachment site.**

amino acids 28-32

**Tyrosine kinase phosphorylation sites.**

amino acids 394-402, 578-585

**N-myristoylation sites.**amino acids 22-28, 27-33, 53-59, 82-88, 162-168, 184-190,  
217-223, 324-330, 325-331, 471-477, 568-574, 759-765**Amidation site.**

amino acids 781-785

**Aminoacyl-transfer RNA synthetases class-II signature 1.**

amino acids 117-138

**Cadherins extracellular repeated domain signature.**

amino acids 121-132, 230-241, 335-346, 439-450, 549-560

483/615

**FIGURE 479**

CTCGGCTGGATTTAAGGTTGCCGCTAGCCGCCTGGGAATTTAAGGGACCCACACTACCTTCCC  
GAAGTTGAAGGCAAGCGGTGATTGTTTGTAGACGGCGCTTTGTCA**ATG**GGACCTGTGCGGTGG  
GAATATTGCTTTTCCTTTTTTTGGCCGTGCACGAGGCTTGGGCTGGGATGTTGAAGGAGGAGG  
ACGATGACACAGAACGCTTGCCCAGCAAATGCGAAGTGTGAAGCTGCTGAGCACAGAGCTAC  
AGGCGGA**ACT**GAGTCGCACCGGT**CGATCTCGAGAGGTGCTGGAGCTGGGGCAGGTGCTGGATA**  
CAGGCAAGAGGAAGAGACACGTGCCTTACAGCGTTTCAGAGACAAGGCTGGAAGAGGCCTTAG  
AGAATTTATGTGAGCGGATCCTGGACTATAGTGTTACAGCTGAGCGCAAGGGCTCACTGAGAT  
ATGCCAAGGGTCAGAGTCAGACCATGGCAACACTGAAAGGCCTAGTGCAGAAGGGGGTGAAGG  
TGGATCTGGGGATCCCTCTGGAGCTTTGGGATGAGCCCAGCGTGGAGGTACATACCTCAAGA  
AGCAGTGTGAGACCATGTTGGAGGAGTTTGAAGACATTGTGGGAGACTGGTACTTCCACCATC  
AGGAGCAGCCCCCTACAAAATTTTCTCTGTGAAGGTCATGTGCTCCAGCTGCTGAAACTGCAT  
GTCTACAGGAACTTGGACTGGAAAGGAGATCACAGATGGGGAAGAGAAAACAGAAGGGGAGG  
AAGAGCAGGAGGAGGAGGAGGAAGAGGAGGAAGAGGAAGGGGGAGACAAGATGACCAAGACAG  
GAAGCCACCCCAA**ACTTG**ACCGAGAAGATCTT**TGA**CCCTTGCCTTTGAGCCCCCAGGAGGGGA  
AGGGATCATGGAGAGCCCTCTAAAGCCTGCACTCTCCCTGCTCCACAGCTTTCAGGGTGTGTT  
TATGAGTGACTCCACCCAAGCTTGTAGCTGTTCTCTCCCATCTAACCTCAGGCAAGATCCTGG  
TGAAACAGCATGACATGGCTTCTGGGGTGGAGGGTGGGGGTGGAGGTCCTGCTCCTAGAGATG  
AACTCTATCCAGCCCCCTAATTGGCAGGTGTATGTGCTGACAGTACTGAAAGCTTTCCTCTTT  
AACTGATCCCACCCCCACCCAAAAGTCAGCAGTGGCACTGGAGCTGTGGGCTTTGGGGAAAGTC  
ACTTAGCTCCTTAAGGTCTGTTTTTAGACCCTTCCAAGGAAGAGGCCAGAACGGACATTCTCT  
GCGATCTATATACATTGCCTGTATCCAGGAGGCTACACACCAGCAAACCGTGAAGGAGAATGG  
GACACTGGGTCATGGCCTGGAGTTGCTGATAATTTAGGTGGGATAGATACTTGGTCTACTTAA  
GCTCAATGTAACCCAGAGCCCACCATATAGTTTTATAGGTGCTCAACTTTCATATATCGCTATT  
AAACTTTTTTCTTTTTTTCTA

484/615

**FIGURE 480**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92256
><subunit 1 of 1, 248 aa, 1 stop
><MW: 28310, pI: 4.63, NX(S/T): 0
MGPVRLGILLFLFLAVHEAWAGMLKEEDDDTERLPSCCEVCKLLSTELQAELSRTGRSREVLE
LGQVLDTGKRKRHPYPSVSETRLEEALENLCERILDYSVHAERKGSRLRYAKGQSQTMATLKGL
VQKGVKVDLGIPLELWDEPSVEVTYLKKQCETMLEEFEDIVGDWYFHHQEQPLQNFLCEGHVL
PAAETACLQETWTGKEITDGEEKTEGEEEEEQEEEEEEEEEGGDKMTKTGSHPKLDREDL
```

**Important features of the protein:****Signal peptide:**

amino acids 1-21

**cAMP- and cGMP-dependent protein kinase phosphorylation site.**

amino acids 106-110

**N-myristoylation site.**

amino acids 115-121

**Amidation site.**

amino acids 70-74

485/615

**FIGURE 481**

GGCGTGTGCAAGGCGGGGTCCGGCCCGCGCAGGTCGGGTAAGCGCGTCTAGGGCGCTGCGCGG  
CGCAGCGAAA**ATG**GCGGCTTCCAGGTGGGCGCGCAAGGCCGTGGTCTGCTTTGTGCCTCTGA  
CCTGCTGCTGCTGCTGCTACTGCTACCACCGCCTGGGTCTGCGCGGCCGAAGGCTCGCCCGG  
GACGCCCCGACGAGTCTACCCACCTCCCCGGAAGAAGAAGGATATTGCGGATTACAATGA  
TGCAGACATGGCGCGTCTTCTGGAGCAATGGGAGAAAGATGATGACATTGAAGAAGGAGATCT  
TCCAGAGCACAAAGAGACCTTCAGCACCTGTCGACTTCTCAAAGATAGACCCAAGCAAGCCTGA  
AAGCATATTGAAAATGACGAAAAAAGGGAAGACTCTCATGATGTTTGTCACTGTATCAGGAAG  
CCCTACTGAGAAGGAGACAGAGGAAATTACGAGCCTCTGGCAGGGCAGCCTTTTCAATGCCAA  
CTATGACGTCCAGAGGTTCAATTGTGGGATCAGACCGTGCTATCTTCATGCTTCGCGATGGGAG  
CTACGCCTGGGAGATCAAGGACTTTTTGGTCGGTCAAGACAGGTGTGCTGATGTAACCTGGA  
GGGCCAGGTGTACCCCGGCAAAGGAGGAGGAAGCAAAGAGAAAAATAAAACAAAGCAAGACAA  
GGGCAAAAAAAGAAGGAAGGAGATCTGAAATCTCGGTCTTCCAAGGAAGAAAATCGAGCTGG  
GAATAAAGAGAAGACCT**GTGA**TGGGGCAGCAGTGACGCGCTGTGGGGGGACAGGTGGACGTG  
GAGAGCTCTTTGCCAGCTCCTGGGGTGGGAGTGGTCTCAGGCAACTGCACACCGGATGACAT  
TCTAGTGTCTTCTAGAAAGGGTCTGCCACATGACCAGTTTGTGGTCAAAGAATTACTGCTTAA  
TAGGCTTCAAGTAAGAAGACAGATGTTTTCTAATTAATACTGGACACTGACAAATTCATGTTT  
ACTATAAAATCTCCTTACATGGAAATGTGACTGTGTTGCTTTTTCCCATTTACACTTGGTGAG  
TCATCAACTCTACTGAGATTCCACTCCCTCCAAGCACCTGCTGTGATTGGGTGGCCTGCTCT  
GATCAGATAGCAAATTCTGATCAGAGAAGACTTTAAACTCTTGACTTAATTGAGTAAACTCT  
TCATGCCATATACATCATTTTCATTATGTTAAAGGTAAAATATGCTTTGTGAACCTCAGATGTC  
TGTAGCCAGGAAGCCAGGGTGTGTAAATCCAAATCTATGCAGGAAATGCGGAGAATAGAAAA  
TATGTCACTTGAAATCCTAAGTAGTTTTGAATTTCTTTGACTTGAATCTTACTCATCAGTAAG  
AGAACTCTTGGTGTCTGTGAGGTTTTATGTGGTCTGTAAAGTTAGGGGTTCTGTTTTGTTTCC  
TTATTTAGGAAAGAGTACTGCTGGTGTGAGGGGTTATATGTTCCATTTAATGTGACAGTTTT  
AAAGGATTTAAGTAGGGAATCAGAGTCCTTTGCAGAGTGTGACAGACGACTCAATAACCTCAT  
TTGTTTCTAAACATTTTCTTTGATAAAGTGCCTAAATCTGTGCTTTTCGTATAGAGTAACATG  
ATGTGCTACTGTTGATGTCTGATTTTGCCGTTTCATGTTAGAGCCTACTGTGAATAAGAGTTAG  
AACATTTATATACAGATGTCATTTCTAAGAACTAAAATTCTTTGGGAAAAACCCTCAAAAAA  
AAAAAAAAAAAAAAAAAAAAA



486/615

**FIGURE 482**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA92289
><subunit 1 of 1, 234 aa, 1 stop
><MW: 26077, pI: 8.13, NX(S/T): 1
MAASRWARKAVVLLCASDLLLLLLLLLPPPGSCAAEGSPGTPDESTPPPRKKKKDIRDYND
ADMARLLEQWEKDDDDIEEGDLPEHKRPSAPVDFSKIDPSKPESILKMTKKGKTLMMFVTV
SGSPTEKETEEITSLWQGSLEFNANYDVQRFIVGSDRAIFMLRDGSYAWWEIKDFLVGQDRC
ADVTLEGQVYPGKGGGSKEKNKTKQDKGKKKKEGDLKSRSSKEENRAGNKREDL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-32

**N-glycosylation site:**

Amino acids 201-205

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 85-89

**Tyrosine kinase phosphorylation site:**

Amino acids 50-59

**N-myristoylation sites:**

Amino acids 30-36;138-144;153-159;176-182

**Amidation site:**

Amino acids 207-211



488/615

**FIGURE 484**

MALPPGPAALRHTLLLLPALLSSGWGELEPQIDGQTWAERALRENERHAFTCRVAGGPGTPRL  
AWYLDGQLQEASTSRLLSVGGEAFSGGTSTFTVTAHRAQHELNCSLQDPRSGRSANASVILNV  
QFKPEIAQVGAKYQEAQGPGLLVVLFALVRANPPANVTWIDQDGPVTVNTSDFLVLDAQNYPW  
LTNHTVQLQLRSLAHNLSVVATNDVGVTASLPAPGPSRHPSLISSDSNNLKLNNVRLPRENM  
SLPSNLQLNDLTPDSRAVKPADRQMAQNNSRPELLDPEPGGLLTSQGFIRLPVLGYIYRVSSV  
SSDEIWL

**N-glycosylation sites:**

amino acids 106-110, 119-123, 162-166, 175-179, 192-196, 205-209,  
251-255, 280-284

**Glycosaminoglycan attachment site:**

amino acids 23-27

**Casein kinase II phosphorylation sites:**

amino acids 36-40, 108-112, 164-168, 282-286, 316-320

**N-myristoylation sites:**

amino acids 34-40, 89-95, 215-221, 292-298, 293-299

489/615

**FIGURE 485**

AGAGTTCCTTTTTCTAGGTCGATTAGGTTATACATTGTTGAAGTATAGTTTCGAGTTAGAATT  
GGTCATTTTATTTTCAGTGTTTTCACAGAAATCGAAGAAGACAGAA**ATG**GCGCTTCTGTGGTGG  
ATATCTACAGTAGCAATACTGTTGTTTACTTCGACGATTTTGGGAACATACGTTGAAGCTGGT  
GCCGCTAAGTCTAACGAAGAAGAGATTGTGAACAAAAGCGAATTTGGAAGATTTCCACGAGGG  
TCGAGAAAGGATGCATCGGGGTGCCACAAGCCGGGCTACCCTGTACCCCTCATTCTCGCTGC  
CCTCCACCTCCCCATGTGCAGCGTCCTCGTCCTATTCTGCATGCT**TAG**TCTAACACCATCAGG  
CTCGTTTATCTTTTCTGTCATTGATCTCACCAGGAGCAAATCACTAGTGCGTGCTTCTGATTTC  
ACGTAACGTAGTATGTAAATAAATGTCAGTGATATTATGAATTGGTAAAACATTTCTGTTATC  
TAAATAAACAGTGAAGTTTGTTTGACTAAAAAAA

490/615

**FIGURE 486**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96855
><subunit 1 of 1, 84 aa, 1 stop
><MW: 9274, pI: 9.70, NX(S/T): 1
MALLWWISTVAILLFTSTILGTYVEAGAAKSNEEEIVNKSEFGRFFPRGSRKDASGCHKPG
YPVPPHSRCPPPPHVQRPRPILHA
```

**Signal peptide:**

Amino acids 1-21

**N-glycosylation site:**

Amino acids 38-42

**N-myristoylation site:**

Amino acids 27-33

491/615

**FIGURE 487**

CGGGGACGGAAGCGGCCCTGGGCCCAGGGGCTGGAGCCGGGCCGGGGCG**ATG**TGGAGCGCG  
GGCCGCGGCGGGGCTGCCTGGCCGGTGCTGTTGGGGCTGCTGCTGGCGCTGTTAGTGCCGGGC  
GGTGGTGCCGCCAAGACCGGTGCGGAGCTCGTGACCTGCGGGTGGTGCTGAAGCTGCTCAAT  
ACGCACCACCGCGTGCGGCTGCACTCGCACGACATCAAATACGGATCCGGCAGCGGCCAGCAA  
TCGGTGACCGGCGTAGAGGCGTCGGACGACGCCAATAGCTACTGGCGGATCCGCGGCGGCTCG  
GAGGGCGGGTGCCCGCGCGGGTCCCCGGTGCGCTGCGGGCAGGCGGTGAGGCTCACGCATGTG  
CTTACGGGCAAGAACCTGCACACGCACCACTTCCCGTCGCCGCTGTCCAACAACCAGGAGGTG  
AGTGCCTTTGGGGAAGACGGCGAGGGCGACGACCTGGACCTATGGACAGTGCGCTGCTCTGGA  
CAGCACTGGGAGCGTGAGGCTGCTGTGCGCTTCCAGCATGTGGGCACCTCTGTGTTCCCTGTCA  
GTCACGGGTGAGCAGTATGGAAGCCCCATCCGTGGGCAGCATGAGGTCCACGGCATGCCCAGT  
GCCAACACGCACAATACGTGGAAGGCCATGGAAGGCATCTTCATCAAGCCTAGTGTGGAGCCC  
TCTGCAGGTCACGATGAACT**TGAG**TGTGTGGATGGATGGGTGGATGGAGGGTGGCAGGTGGG  
GCGTCTGCAGGGCCACTCTTGGCAGAGACTTTGGGTTTGTAGGGGTCCTCAAGTGCCTTTGTG  
ATTAAAGAATGTTGGTCTATGAAA

492/615

**FIGURE 488**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96857  
><subunit 1 of 1, 221 aa, 1 stop  
><MW: 23598, pI: 6.96, NX(S/T): 0  
MWSAGRGGAAWPVLLGLLLALLVPGGGAAKTGAELVTCGSVLKLLNTHHRVRLHSHDIKYGSG  
SGQQSVTGVEASDDANSYWRIRGGSEGGCPRGSPVRCGQAVRLTHVLTGKNLHTHHFPSPLSN  
NQEVSAFGEDGEGDDLDLWTVRCSGQHWEREAAVRFQHVGTSVFLSVTGEQYGSPIRGQHEVH  
GMPSANTHNTWKAMEGIFIKPSVEPSAGHDEL

**Important features of the protein:****Signal peptide:**

amino acids 1-28

**Glycosaminoglycan attachment site.**

amino acids 62-66

**N-myristoylation sites.**amino acids 16-22, 25-31, 27-33, 61-67, 71-77, 86-92, 87-93,  
91-97, 190-196**Endoplasmic reticulum targeting sequence.**

amino acids 218-223

493/615

**FIGURE 489**

CAGCAGCCGAGACAGCAGCTGAGACGGCAGCGGCAGCTTCTCAGGGCCGGAGCCAGTTCTTGGAGGAGACTCTGC  
ACAGGGGCATGGGATCACTGTGGTGCCCTTTTCTGTGCCTGTGCCTTCTGACTTTGCAGAATGCAACAACAGAGAC  
ATGGGAAGAACTCCTGAGCTACATGGAGAATATGCAGGTGTCCAGGGGCCGGAGCTCAGTTTTTCTCTCGTCA  
ACTCCACCAGCTGGAGCAGATGCTACTGAACACCAGCTTCCCAGGCTACAACCTGACCTTGACAGACCCACCAT  
CCAGTCTCTGGCCTTCAAGCTGAGCTGTGACTTCTCTGGCCTCTCGCTGACCAGTGCCACTCTGAAGCGGGTGCC  
CCAGGCAGGAGGTGAGCATGCCCGGGGTGAGCAGGCCATGCAGTTCCCCGCCGAGCTGACCCGGGACGCCTGCAA  
GACCCGCCCCAGGGAGCTGCGGCTCATCTGTATCTACTTCTCCAACACCCACTTTTTCAAGGATGAAAACAACCTC  
ATCTCTGCTGAATAACTACGTCTGGGGGCCAGCTGAGTCATGGGCACGTGAACAACCTCAGGGATCCTGTGAA  
CATCAGCTTCTGGCACAACCAAGCCTGGAAGGCTACACCTGACCTGTGTCTTCTGGAAGGAGGGAGCCAGGAA  
ACAGCCCTGGGGGGGCTGGAGCCCTGAGGGCTGTGCTACAGAGCAGCCCTCCCCTCTCAGGTGCTCTGCCGCTG  
CAACCACTCACCTACTTTGCTGTTCTCATGCAACTCTCCCAGCCCTGGTCCCTGCAGAGTTGCTGGCACCTCT  
TACGTACATCTCCCTCGTGGGCTGCAGCATCTCCATCGTGGCCTCGCTGATCAGAGTCTGCTGCACTTCCATTT  
CAGGAAGCAGAGTGACTCCTTAACACGTATCCACATGAACCTGCATGCCTCCGTGCTGCTCCTGAACATCGCCTT  
CCTGCTGAGCCCCGCATTGCAATGTCTCCTGTGCCCGGGTGCAGCATGCACGGCTCTGGCCGCTGCCCTGCACTA  
CGCGCTGCTCAGCTGCCTCACCTGGATGGCCATCGAGGGCTTCAACCTCTACCTCCTCCTCGGGCGTGTCTACAA  
CATCTACATCCGCAGATATGTGTTCAAGCTTGGTGTGCTAGGCTGGGGGGCCCCAGCCCTCCTGGTGTGCTTTT  
CCTCTCTGTCAAGAGCTCGGTATACGGACCCTGCACAATCCCCGTCTTCGACAGCTGGGAGAATGGCACAGGCTT  
CCAGAACATGTCCATATGCTGGGTGCGGAGCCCCGTGGTGCACAGTGTCTGGTCATGGGCTACGGCGGCCCTCAC  
GTCCCTCTTCAACCTGGTGGTGTGCTGGCCTGGGCGCTGTGGACCCTGCGCAGGCTGCGGGAGCGGGCGGATGCACC  
AAGTGTGAGGGCCTGCCATGACACTGTCACTGTGCTGGGCCTCACCGTGTGCTGGGAACACCTGGGCCTTGGC  
CTTCTTTTCTTTTGGCGTCTTCTCTGCTGCCCCAGCTGTTCTCTTCAACCATCTTAAACTCGCTGTACGGTTTCTT  
CCTTTTCTGTGGTTCTGCTCCCAGCGGTGCCGCTCAGAAGCAGAGGCCAAGGCACAGATAGAGGCCTTCAGCTC  
CTCCCAACAACACAGTAGTCCGGGCCTCCTGGCCTGGAATCCTCAGCCTCTCTGCCCCCAGTAGGCTGAGGCT  
ACGGCTCCTGCTAGAGAGGGTGGCAGGCCTGCTGCTGGACCCAGAGGCCACTGTGACCGCCAAGGGGCCTTTTTC  
CACTTCCACGGCCTCTCCAGGCACTGAGGGGAAGGCATTGCTCTACCTCTCCCTGACATTTTGCTCCGGGGCAGA  
TCCAACCTTACCTGGGGCAGCAAACTTTGTCTGTACCTGGGCCAGCTCGCCAGGGATGTGGGCAGAGCACC  
GCCTGGGCATCAGGAAGCCAAGTTTCAAGGACTGTCTTTGAGTCTGTCTGTATGACCTTGGGCCTGCCACTTCTC  
ACAGACCCTAGGTATCCACAGCTGTGACATGGGGGCAAGCAGCTTTGTTTCAGCCTAACCCAGGAGCTTAGTAAA  
AATTGCATAAGACCAGGGGAAGAGTGTGACGTGGGGTGGGAATTCCCGCGGCCTCCACCTGCTTGGTAGGGGC  
AGGATCTCATTAGGCTGCCCTGGAAGCACCTGCTTGGCCCTGCCACCTTCTCCAGGGGAGGGCCAGATGGCAT  
CCTGGCTTGGGGCGGGTGGGACCTACCCAGGCTCTGAGACTTTACTGGCCTATGCCTGAGGCCTCTTTTCTTTTA  
ACTCCCTAAATTATGATGACTCCAAGTCCAAGCCACCCTTCCCAAAGATTGGGAGGTTCCGCCGTTCCAGAGG  
CTCCTCCTGCGGTGCTCCCAAGACTTCCATAGACCATCTGGACCAGTAGCCCATCCCGCAGTTTTCTTGGGGGCA  
GAGGAAAACGCTTCTTTCTCCTCCAGCTGAATCAGCTGGATCCCAGTGTCTGGCTGTTTGGTGATTGGGCAAGA  
TTGAATTTGCCCAGGTAGGCGTGAGAGTGTGGGTTTTAAATTCGAAGCTCAGGCCATAGTTTCAGAGAATCACCC  
TTACCCCAAGACCTTTCATGAGACAGTGTCTCATGAAGCCAGTGCGTTTCCAGAACGAACACTAGGCGGCACCGTTG  
GTCCACACTCAGAGGCCCTTGGCGCCAAGACTGCATCTAGAATCGCTCAAACACCTGTTTGCAGACCCCATGCAC  
CAGCTGGAGGGGCCGTAACCTGCAGGACTGCGCCTACTGAGTGACCCATTTCTCCAGGAGGAAAGGCAAGACAG  
CTTACACGGCCATTTGTCTCTTTTCCCAATGCGGCGGTGCACTTTGCTCTTGGGGGCTGCACCCAGACATAGC  
TGGCACCAGAGCAGGGTGTGAGGTGGTGGTGTGCTCAGGGCCCTGCCCCAGGCCACTGGGCGTTTTGATGACCT  
CAAAGGTACAGGCAGAAAATAGGAGCAGGATTTCCCTGGGGAAAAGTTATCCTGGGACATCTTCTGCTCTTCT  
GTACATTTCTAGATGCAATAACTCCTTCACCAAGCAGTGAGTGGCGTAGGCTCTGGAGCCAGGCTGCCTGGGCT  
CCAATGCCAGCTCTGCCACTTGTAGCTGTGAGACTGTGGACAAACCACTCAGCCTCTGTGTGCTCAGTTTTCC  
TATTTGTAATAATAGAGACCATAGTGGTACCTATTTTGAAGACTAAGTAAAAGAATTCAAATAAAGAGACTTGGCA  
CAGAGTAAGTGCTCAGTAAAAA



494/615

**FIGURE 490**

```

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96860
><subunit 1 of 1, 528 aa, 1 stop
><MW: 59000, pI: 8.73, NX(S/T): 9
MDHCGALFLCLCLLTQONATTETWEELLSYEMNQVSRGRSSVFSSRQLHQLEQMLLNTS
FPGYNLTQTPTIQSLAFKLSCDFSGLSLTSATLKRVPQAGGQHARGQHAMQFPAELTRD
ACKTRPRELRLICIFYFSNTHFFKDENNSSLNNYVLGAQLSHGHVNNLRDPVNI SFWHNQ
SLEGYTLTCVFWKEGARKQPWGGWSPEGCRTEQPSHSQVLCRCNHLTYFAVLMQLSPALV
PAELLAPLTYISLVGCSISIVASLITVLLHFHFRKQSDSLTRIHMNLHASVLLLNIAFLL
SPAFAMSPVPGSACTALAAALHYALLSCLTWMAIEGFNLYLLLGRVYNIYIRRYVFKLG
LGWGAPALLVLLSLSVKSSVYGPCTIPVFDSEWNGTGFQNMSCWVRSPVHSLVLMGYG
GLTSLFNLVVLAWALWTLRRLRERADAPSVRACHDTVTVLGLTVLLGTTWALAFFSFGVF
LLPQLFLFTILNSLYGFFLFLWFCSQRCRSEAEAKAQIEAFSSSQTTQ

```

**Important features of the protein:****Signal peptide:**

Amino acids 1-21

**Transmembrane domains:**

Amino acids 244-264;290-309;316-344;358-376;411-431;468-491

**N-glycosylation sites:**Amino acids 18-22;58-62;65-69;146-150;147-151;173-177;  
179-183;394-398;400-404**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 274-278

**N-myristoylation sites:**

86 GLSLTS  
 101 GGQHAR  
 157 GAQLSH  
 255 GCSISI  
 311 GSACTA  
 420 GGLTSL  
 467 GTTWAL

**Prokaryotic membrane lipoprotein lipid attachment sites:**

Amino acids 246-257;318-329

**Eukaryotic thiol (cysteine) proteases histidine active site:**

Amino acids 410-421

**G-protein coupled receptors family 2 proteins:**

Amino acids 273-302;314-343

495/615

**FIGURE 491**

CTTGGCTGCCCCGACAACAAGCTCGCCACCTGCGCTGGGCGCATCCACCATCCAAGGCCAGCT  
GAGGGGCACCAGACAGAGG**ATG**AGGAGAGAGAGTCCGACACGGGCTGCCCTGAGAGACATTTT  
CATGGACATCCTCATGCTGCTTCTGCTTTTGTGTGTAATATATGGGAGATTTTCCCAAGATGA  
ATACTCCCTCAATCAAGCTATCCGGAAAGAATTTACAAGAAATGCCAGAACTGCTTGGGTGG  
CCTGAGAAACATCGCTGACTGGTGGGACTGGAGTCTGACCACACTTCTGGATGGCCTGTACCC  
GGGAGGCACCCCGTCAGCCCGTGTGCCGGGGGCTCAGCCTGGAGCTCTTGGAGGAAAATGCTA  
CCTAATAGGCAGTTCCGTAATTAGGCAGCTAAAAGTTTTTCTAGGCATTTATGCAAGCCTCC  
CAGGCCATTTTTCAGCACTCATCGAAGACTCTATTCTTACATGTAGTCCCGAAGTTGGAGGCCC  
TGAGAACCCCTACCTGATAGACCCAGAGAACCAAAACGTGACCCTGAATGGTCCTGGGGGCTG  
TGGGACAAGGGAGGACTGTGTGCTCAGCCTGGGCAGAACAAAGGACTGAAGCCCACACAGCCCT  
GTCCCGACTCAGGGCCAGCATGTGGATTGACCGCAGCACCAGGGCTGTGTCTGTGCACTTCAC  
TCTCTATAACCCTCCAACCCAACTCTTCACCAGCGTGTCCCTGAGAGTGGAGATCCTCCCTAC  
GGGGAGTCTCGTCCCCTCATCCCTGGTGGAGTCATTGAGCATCTTCCGCAGCGACTCAGCCCT  
GCAGTACCACCTCATGCTTCCCCAGCTGGTCTTCTTGGCACTCAGCCTGATCCACCTCTGTGT  
TCAACTCTACCGTATGATGGACAAGGGCGTCCTCAGCTACTGGCGAAAGCCAAGGAACTGGCT  
GGAGGTAGCCTCTCTTGTGTGTCATTTTCTTTTGAAAA**TA**ACAATAAACTGTTTATATCTTGAA  
AAAATAATTTAAATAAGAAATTGATTATGCACTAGCTACTGCCAACATTATTGCAGTTTTCTC  
CCTCTGTAGTGTTAATCTCAAAACAGCATTTGAGATCAGGTATCATTTAGTGTTGTACAGTT  
ACCGTCATGTACCACACGAATTTAGCCAAAGGTGGTGGTCCATAAGATCATATGGTGCTAAG  
AAATTTCTGTACCTAATGACATCTTGATTCTGACCTTGTATGTAGGCCTAGGCTAAATATGT  
CTGTTTGTATCTTAGCTTTTAATAAAGAAGTTTAAAAATAAAAAA

496/615

**FIGURE 492**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96861
><subunit 1 of 1, 300 aa, 1 stop
><MW: 33649, pI: 9.26, NX(S/T): 1
MRRESRTRAALRDISMDILMLLLLCVIYGRFSQDEYSLNQAIRKEFTRNARNCLGGLRN
IADWWDWSLTLLDGLYPGGTPSARVPGAQPGALGGKCYLIGSSVIRQLKVFPRLCKPP
RPFSAIEDSIPTCSPEVGGPENPYLIDPENQNVTLNGPGGCGTREDCVLSLGRTRTEAH
TALSRLRASMWIDRSTRAVSVHFTLYNPPTQLFTSVSLRVEILPTGSLVPSSLVESFSIF
RSDSALQYHMLPQLVFLALSLIHL CVQLYRMDKGVLSYWRKPRNWLEVASLVSFSEK
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-30

**Transmembrane domain:**

Amino acids 250-267

**N-glycosylation site:**

Amino acids 153-157

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 2-6

**N-myristoylation sites:**

Amino acids 56-62;75-81;79-85;80-86;88-94;92-98;160-166

497/615

**FIGURE 493**

TCTCAGGGCTTCATACAGGAAATCTATTGCTGTGTCAAGTTCCAGAGAAAAGCTTCTGTTCGT  
CCAAGTTACTAACCAGGCTAAACCACATAGACGTGAAGGAAGGGGCTAGAAGGAAGGGAGTGC  
CCCCTGTTGATGGGGTAAGAGGATCCTGTACTGAGAAGTTGACCAGAGAGGGTCTCACCATG  
CGCACAGTTCCTTCTGTACCTGTGTGGAGGAAAAGTACTGAGTGAAGGGCAGAAAAAGAGAAA  
ACAGAAATGCTCTGCCCTTGGAGAACTGCTAACCTAGGGCTACTGTTGATTTTGACTATCTTC  
TTAGTGGCCGAAGCGGAGGGTGCTGCTCAACCAAACAACCTCATTAATGCTGCAAACTAGCAAG  
GAGAATCATGCTTTAGCTTCAAGCAGTTTATGTATGGATGAAAAACAGATTACACAGAACTAC  
TCGAAAGTACTCGCAGAAGTTAACACTTCATGGCCTGTAAAGATGGCTACAAATGCTGTGCTT  
TGTTGCCCTCCTATCGCATTAAGAAATTTGATCATAATAACATGGGAAATAATCCTGAGAGGC  
CAGCCTTCCTGCACAAAAGCCTACAGGAAAGAAACAAATGAGACCAAGGAAACCAACTGTACT  
GATGAGAGAATAACCTGGGTCTCCAGACCTGATCAGAATTCCGACCTTCAGATTTCGTCCAGTG  
GCCATCACTCATGACGGGTATTACAGATGCATAATGGTAACACCTGATGGGAATTTCCATCGT  
GGATATCACCTCCAAGTGTTAGTTACACCTGAACTGACCCTGTTTCAAAACAGGAATAGAAT  
GCAGTATGCAAGGCAGTTGCAGGGAAGCCAGCTGCGCAGATCTCCTGGATCCCAGAGGGCGAT  
TGTGCCACTAAGCAAGAATACTGGAGCAATGGCACAGTGAAGTGTAAAGAGTACATGCCACTGG  
GAGGTCCACAATGTGTCTACCGTGACCTGCCACGTCTCCCATTTGACTGGCAACAAGAGTCTG  
TACATAGAGCTACTTCCTGTTCCAGGTGCCAAAAAATCAGCAAAATTATATATTTCCATATATC  
ATCCTTACTATTATTATTTTGACCATCGTGGGATTCATTTGGTTGTTGAAAGTCAATGGCTGC  
AGAAAAATATAAATTGAATAAAACAGAATCTACTCCAGTTGTTGAGGAGGATGAAATGCAGCCC  
TATGCCAGCTACACAGAGAAGAACAATCCTCTCTATGATACTACAAACAAGGTGAAGGCATCT  
CAGGCATTACAAAGTGAAGTTGACACAGACCTCCATACTTTATAAGTTGTTGGACTCTAGTAC  
CAAGAAACAACAACAAACGAGATACATTATAATTACTGTCTGATTTTCTTACAGTTCTAGAAT  
GAAGACTTATATTGAAATTAGGTTTTCCAAGGTTCTTAGAAGACATTTTAATGGATTCTCATT  
CATACCCTTGTATAAATTGGAATTTTTGATTCTTAGCTGCTACCAGCTAGTTCTCTGAAGAACT  
GATGTTATTACAAAGAAAATACATGCCCATGACCAAATATTCAAATTGTGCAGGACAGTAAAT  
AATGAAAACCAATTTCCCTCAAGAAATAACTGAAGAAGGAGCAAGTGTGAACAGTTTCTTGTTG  
TATCCTTTCAGAATATTTTAATGTACATATGACATGTGTATATGCCTATGGTATATGTGTCAA  
TTTATGTGTCCCTTACATATACATGCACATATCTTTGTCAAGGCACCAGTGGGAACAATACA  
CTGCATTACTGTTCTATACATATGAAAACCTAATAATATAAGTCTTAGAGATCATTTTATATC  
ATGACAAGTAGAGCTACCTCATTCTTTTTAATGGTTATATAAAATTCCATTGTATAGTTATAT  
CATTATTTAATTAAAAACAACCCTAATGATGGATATTTAGATTCTTTTAAGTTTTGTTTATTT  
CTTTTAAGTTTTGTTTGTGGTATAAACAATACCACATAGAATGTTTCTTGTTTATATATCTCT  
TTGTTTTTGAGTATATCTGTAGGATAACTTTCTTGAGTGGAAATTGTCAGGTCAAAGGGTTTTGT  
GCATTTTACTATTGATATATATGTTAAATTGTGTCAAATATATATGTCAAATTCCTCCAACA  
TTGTTTTAATGTGCCTTTCCCTAAATTTCTATTTTAATAACTGTACTATTCTCTCTACAG  
TTGCCACTTTCTCTTTTAAATCAACCAGATTAAATATGATGTGAGATTATAATAAGAATTATA  
CTATTTAATAAAAATGGATTTATA

498/615

**FIGURE 494**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96866
><subunit 1 of 1, 348 aa, 1 stop
><MW: 39069, pI: 8.13; NX(S/T): 10
MLCPWRTANLGLLLIILTIFLVAEAEGAAQPNNSLMLQTSKENHALASSSLCMDEKQITQN
YSKVLAEVNTSWPVKMATNAVLCCPPIALRNLIITWEIILRGQPSCTKAYRKETNETKE
TNCTDERITWVSRPDQNSDLQIRPVAITHDGYRCIMVTPDGNFHRGYHLQVLVTPELTL
FQNRNRTAVCKAVAGKPAAQISWIPEGDCATKQEYWSNGTVTVKSTCHWEVHNVSTVTCH
VSHLTGNKSLYIELLPVPGAKKSAPLYIPYIILTIILTIIVGFIWLLKVNGCRKYKLNKT
ESTPVVEEDEMOPYASYTEKNNPLYDTTNKVKASQALQSEVDTDLHTL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-24

**Transmembrane domains:**

Amino acids 78-98;267-286

**N-glycosylation sites:**Amino acids 31-35;60-64;69-73;116-120;122-126;185-189;  
218-222;233-237;247-251;298-302**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 112-116

**N-myristoylation sites:**

Amino acids 103-109;259-265

499/615

**FIGURE 495**

CCAGGTGCACAGCGCATCGCCCGAGGCTGTCACCGCCCTGCCCCGCCCACCCCAGCTGTCCTG  
GACCCAGGGGCAGGGAGAGGCTGGACGCCAGGTGCGCGGACACAGAAGCGTCTAAGCACAGCT  
TCCTCCTTGCCGCTCCGGGAAGTGGGCAGCCAGCCCAGGAACCAGTACCACCTGCACC**ATGGG**  
GCTGTCCCGGAAGGAGCAGGTCTTCTTGCCCTGCTGGGGGCTCGGGGGTCTCAGGCCTCAC  
GGCACTCATTCTCCTCCTGGTGGAGGCCACCAGCGTGCTCCTGCCACAGACATCAAGTTTGG  
GATCGTGTGATGCGGGCTCCTCCACACGTCCCTCTTCTGTATCAGTGGCCGGCGAACAA  
GGAGAATGGCACGGGTGTGGTCAGCCAGGCCCTGGCCTGCCAGGTGGAAGGGCCTGGAATCTC  
CTCCTACACTTCTAATGCTGCACAGGCTGGTGAGAGCCTGCAGGGCTGCTTGGAGGAGGCGCT  
GGTGCTGATCCAGAGGGCCAGCATCGGAAAACACCCACGTTCTTGGGGGCCACGGCTGGCAT  
GAGGTTGCTCAGCCGGAAGAACAGCTCTCAGGCCAGGGACATCTTTGCAGCAGTCACCCAGGT  
CCTGGGGCCGGTCTCCCGTGGACTTTTGGGGTGCCGAGCTCCTGGCCGGGCAGGCCGAAGGTGC  
CTTTGGTTGGATCACTGTCAACTACGGCTTGGGGACGCTGGTCAAGTACTCCTTCACTGGAGA  
ATGGATCCAGCCTCCGGAGGAGATGCTGGTGGGTGCCCTGGACATGGGAGGGGCTCCACCCA  
GATCACGTTCTGTCCTGGGGGCCCCATCTTGGACAAGAGCACCCAGGCCGATTTTCGCCTCTA  
CGGCTCCGACTACAGCGTCTACACTCACAGCTACCTGTGCTTTGGACGGGACCAGATGCTGAG  
CAGGCTCCTCGTGGGGCTGGTGCAGAGCCGCCCGCTGCCCTGCTCCGTCACCCGTGCTACCT  
CAGCGGCTACCAGACCACACTGGCCCTGGGCCCGCTGTATGAGTCACCCCTGTGTCCACGCCAC  
GCCCCCGCTGAGCCTCCCCCAGAACCTCACAGTTGAAGGGACAGGCAACCCTGGAGCCTGCGT  
CTCAGCCATCCGGGAACTTTTCAACTTCTCCAGCTGCCAGGGCCAGGAGGACTGCGCCTTTGA  
CGGGGTCTACCAGCCCCCGCTGCGGGGCCAGTTCTATGTGGAGGCCAGCTACCCTGGGCAGGA  
CCGCTGGCTGCGGGACTACTGTGCCTCAGGCCTGTACATCCTCACCCCTCCTGCACGAGGGCTAC  
GGTTTCAGCGAGGAGACCTGGCCCAGCCTCGAGTTCGGAAGCAGGCGGGCGGTGTGGACATT  
GGCTGGACACTGGGCTACATGCTGAACCTGACCGGGATGATCCCGGCCGATGCGCCGGCTCAG  
TGGCGGGCAGAGAGCTACGGCGTCTGGGTGGCCAAAGTGGTGTTCATGGTGCTGGCCCTGGTG  
GCGGTGGTGGGGGCTGCCTTGGTCCAGCTCTTCTGTTGCAGGACT**AGT**GGGAAGGCGGAGGT  
GGGCCCCCACAGAGCCCACAGGCAGCTGCGTCCCGGATGCTGGAGGCTTCTGAGCCCTGAGC  
GCCGTGGGGCCTTGCTCTGTGGCTCTGCCACGGTCAGGTGACAGCCACCTCCAGGGCACCGT  
CAGGGTGGTGCTGGCCACAGAGGCTGCATGACCTCCCCTCCCGGCGTCCCTCCCCAACCTCC  
TTCCGCAACTGGGCTTCCAGGGCCGTAGGTGCCTTTCTGCACACAGGCCGCCAGGACTCGTGG  
TGTCTCCAGGCTGTGTGACTGCAGGGCCACATGCTGCCTGCAAACAGGGCAAGACCACGGAGG  
CACAGGGGTCTGCTCCTGATGGGGCCTCAGGAGGGGCGGAGAGGGGTGGAAGGGAGGGAGCT  
GCCCCACCTGGACCCCCGCTCTCCCTGCTGTTGTCTGAGCAGATGGATGGAGTCCAGGCCTGG  
GGGCTTCTGCTGGGCCAGCCCGCCTCCACACCCACTTGGAGGGTGAGACTGCAGTGGGGGT  
TGTTTTTATTAAAAGCATCATGGACACAGCAAAAAAAAAAAAAAAAAA

500/615

**FIGURE 496**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96870
><subunit 1 of 1, 458 aa, 1' stop
><MW: 49377, pI: 4.98, NX(S/T): 5
MGLSRKEQVFLALLGASGVSGLTALILLVLEATSVLLPTDIKFGIVFDAGSSHTSLFLYQ
WPANKENGTGVVSQALACQVEGPGISSYTSNAAQAGESLQGCLEEALVLIPEAQHRKTPT
FLGATAGMRLLSRKNSSQARDIFAAVTQVLGRSPVDFWGAELLAGQAEGAFGWITVNYGL
GTLVKYSFTGEWIQPPPEEMLVGALDMGGASTQITFVPGGPILDKSTQADFRLYGSDYSVY
THSYLCFGRDQMLSRLLVGLVQSRPAALLRHPCYLSGYQTTLALGPLYESPCVHATPPLS
LPQNLTVEGTGNPGACVSAIRELENFSSCQGQEDCAFDGVYQPPLRGQFYVEASYPGQDR
WLRDYCASGLYILTLLHEGYGFSEETWPSLEFRKQAGGVDIGWTLGYMLNLTGMIPADAP
AQWRAESYGVVWAKVVFVVLALVAVVGAALVQLFWLQD
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-21

**Transmembrane domain:**

Amino acids 428-449

**N-glycosylation sites:**

Amino acids 67-71;135-139;304-308;325-329;410-414

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 133-137

**N-myristoylation sites:**Amino acids 50-56;123-127;165-171;207-213;234-240;  
259-265;311-317;314-320;331-337;398-404;  
413-419;429-435**GDA1/CD39 family of nucleoside phosphatases proteins:**

Amino acids 43-59;202-215

501/615

**FIGURE 497**

GCCTTATAAAGTAGCCTCTGCATCTGCCTGCCTCGGGCAGAGGAGGGCTACCCTGGGGCTGAG  
AGTTCACCTGTCTCAGGAACCACCTGAGCCACAGATCCTGTGGGCAGCGGCCAGGGCAGCCA  
**TGG**CCTTGGGCAAGTAGGCTGGGCCTGCTGCTGGCACTGCTGCTGCCCCGTGGTGGTGCCTCCA  
CGCCAGGCACCGTGGTCCGACTCAACAAGGCAGCATTGAGCTACGTGTCTGAAATTGGGAAAG  
CCCCCTCTCAGCGGGGCCCTGCAGGTCACTGTCCCTCATTTCTGGACTGGAGTGGAGAGGCGC  
TTCAGCCCACCAGGATCCGGATTCTGAATGTCCATGTGCCCCGCTCCACCTGAAATTCATTG  
CTGGTTTCGGAGTGCGCCTGCTGGCAGCAGCTAATTTTACTTTCAAGGTCTTTCGCGCCCCAG  
AGCCCCCTGGAGCTGACGCTGCCTGTGGAAGTGTGGCTGACACCCGCGTGACCCAGAGCTCCA  
TCAGGACCCCTGTGGTCAGCATCTCTGCCTGCTCTTTATTCTCGGGCCACGCCAACGAGTTTG  
ATGGCAGTAACAGCACCTCCACGCGCTGCTGGTCTGGTGCAGAAGCACATTAAAGCTGTCT  
TGAGTAACAAGCTGTGCCTGAGCATCTCCAACCTGGTGCAGGGTGTCAATGTCCACCTGGGCA  
CCTTAATTGGCCTCAACCCCGTGGGTCTGAGTCCAGATCCGCTATTCCATGGTCACTGTGC  
CCACTGTCACCAGTGACTACATTTCCCTGGAAGTCAATGCTGTTCTCTTCCCTGCTGGGCAACC  
CCATCATCCTGCCCCACGGATGCCACCCCTTTTGTGTTGCCAAGGCATGTGGGTACCGAGGGCT  
CCATGGCCACCGTGGGCCTCTCCAGCAGCTGTTTGACTCTGCGCTCCTGCTGCTGCAGAAGG  
CCGGTGGCCTCAACCTGGACATCACAGGGCAGCTGAGGTGGATGACAACCTGCTGAACACCT  
CTGCTCTGGGCCGGCTCATCCCGGAGGTGGCCCGCCAGTTTCCCGAGCCCATGCCTGTGGTGC  
TCAAGGTGCGGCTGGGTGCCACACCTGTGGCCATGCTCCACACAAACAACGCCACCCCTGCGGC  
TGCAGCCCTTCGTGGAGGTCTGGCCACAGCCTCCAACCTCGGCTTTCAGTCCCTCTTCTCCC  
TGGATGTGGTAGTGAAGTTGAGACTCCAGCTCTCTGTGTCCAAGGTGAAGCTTCAGGGGACCA  
CGTCTGTGCTGGGGGATGTCCAGCTCACGGTGGCCTCCTCCAACGTGGGCTTCATTGATACAGAT  
CAGGTGCGCACACTGATGGGCACCGTTTTTGAGAAGCCCCTGCTGGACCATCTCAATGCTCTC  
TTGGCCATGGGAATTGCCCTCCCTGGTGTGGTCAACCTCCACTATGTTGCCCTGAGATCTTT  
GTCTATGAGGGCTACGTGGTGATATCCAGTGGACTCTTCTACCAGAGCT**TGA**GGCAAGACCACT  
GGGAGGCCTGAGAGTGGGCCAGCTCGCTGCTCAGGCGAATTTCTCATTTCAAGCCACTGGGGA  
AACTGAGGCAAAACCATACTTAGTCATCACCAACAAGCTGGACTGCTTAGCTGGGCTGTTTTA  
TCTTCCCTGAGTGCCTGGGTCTCCCTCCCTCACTTCTGCCCTTCCCTTCCCTCCTCCTCTTCT  
CCTCCCTCTTCCCTCATCTCCCCCTCCTTCTGCCCCACCCAGGGGGGAGCAGACTGCT  
CCTCCAGGCTGTATAGACCTGCCCTCTTGCAATTAAACAAGTTCTCTTGAGCTGC



502/615

**FIGURE 498**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96872
><subunit 1 of 1, 458 aa, 1 stop
><MW: 49158, pI: 8.72, NX(S/T): 4
MAWASRLGLLLALLLPVVGASTPGTVVRLNKAALSYVSEIGKAPLQRALQVTVPHFLDWS
GEALQPTRIRILNVHVPRLHLKFIAGFGVRLLAANFTFKVFRAPEPLELTLPVELLADT
RVTQSSIRTPVVSISACSLFSGHANEFDGSNSTSHALLVLVQKHIKAVLSNKLCLSISNL
VQGVNVHLGTLIGLNPVGPESQIRYSMVSVPTVTSDYISLEVNAVLFLLGNPIILPTDAT
PFVLPRHVGTEGSMATVGLSQQLEFDSALLLLQKAGALNLDITGQLRSDDNLLNTSALGRL
IPEVARQFPEPMPVVLKVR LGATPVAMLHTN NATLRLQPFVEVLATASNSAFQSLFSLDV
VVNLRLQLSVSKVKLQGTTSVLGDVQLTVASSNVGFIDTDQVRTLMGTVF EKPLL DHLNA
LLAMGIALPGVVNLHYVAPEIFVYEGYVVISSGLFYQS
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**Transmembrane domain:**

Amino acids 217-236

**N-glycosylation sites:**

Amino acids 96-100;151-155;293-297;332-336

**N-myristoylation sites:**

Amino acids 8-14;149-155;189-195;249-255;252-258;283-289

**LBP / BPI / CETP family proteins:**

Amino acids 22-50; 251-287

503/615

**FIGURE 499**

TTGAAAATCTACTCTATCAGCTGCTGTGGTTGCCACCATTCTCAGGACCCTCGCC**ATG**GAAAGC  
CCTTATGCTGCTCACCCCTGTCTGTTCTGCTCTGCTGGGTCTCAGCTGACATTGCTGTCACTC  
CTGCTACAAGGTCCCTGTGCTGGGCTGTGTGGACCGGCAGTCCTGCCGCCTGGAGCCAGGACA  
GCAATGCCTGACAACACATGCATACCTTGGTAAGATGTGGGTTTTCTCCAATCTGCGCTGTGG  
CACACCAGAAGAGCCCTGTGAGGAGGCCTTCAACCAAAACCAACCGCAAGCTGGGTCTGACATA  
TAACACCACCTGCTGCAACAAGGACAACCTGCAACAGCGCAGGACCCCGGCCCACTCCAGCCCT  
GGGCCTTGTCTTCCTTACCTCCTTGGCTGGCCTTGGCCTCTGGCTGCTGCAC**TGA**ACTCATT  
CCATTGGCTGCCCCCTCCTCCACCTGCCTTGGCCTGAGCCTCTCTCCCTGTGTCTCTGTATCC  
CCTGGCTTTACAGAATCGTCTCTCCCTAGCTCCCATTTCTTTAATTAAACACTGTTCCGAGTG  
GTCTCCTCATCCATCCTTCCACCTCACACCCTTCACTCTCCTTTTTCTGGGTCCCTTCCAC  
TTCCTTCCAGGACCTCCATTGGCTCCTAGAAGGGCTCCCCACTTTGCTTCCTATACTCTGCTG  
TCCCCTACTTGAGGAGGGATTGGGATCTGGGCCTGAAATGGGGCTTCTGTGTTGTCCCCAGTG  
AAGGCTCCCACAAGGACCTGATGACCTCACTGTACAGAGCTGACTCCCCAAACCCAGGCTCCC  
ATATGTACCCCATCCCCATACTCACCTCTTCCATTTTGAGTAATAAATGTCTGAGTCTGGA  
AAAAAAAAAAAAAAAAAAAA

504/615

**FIGURE 500**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96878
><subunit 1 of 1, 125 aa, 1 stop
><MW: 13821, pI: 8.60, NX(S/T): 2
MKALMLLTLSVLLCWVSADIRCHSCYKVPVLGCVDRQSCRLEPGQQCLTTHAYLGKMWVFSNL
RCGTPEEPCQEAFNQTNRKLGLTYNNTCCNKDNCNSAGPRPTPALGLVFLTSLAGLGLWLLH
```

**Important features of the protein:****Signal peptide:**

amino acids 1-18

**N-glycosylation sites.**

amino acids 77-81, 88-92

**N-myristoylation site.**

amino acids 84-90

**Ly-6 / u-PAR domain protein signature.**

amino acids 85-98

505/615

**FIGURE 501**

GGAGCCTCCTAATGCAGTCTTCTGCACAGTCCTGGGGACTGACTGACTGAATCACACCTCTGG  
GGCTGGGGGCTGCTGACATGTGTGCCTTTCCTTGGCTGCTTCTTCTCCTGCTGCTCCAGGAGG  
GCAGCCAAAGGAGACTCTGGAGATGGTGTGGATCCGAGGAAGTGGTTGCGGTCTTCAGGAGT  
CCATCAGCCTCCCCCTGGAAATACCACCAGATGAAGAGGTTGAGAACATCATCTGGTCCTCTCAC  
AAAAGTCTTGCCACTGTGGTGCCAGGGAAAGAGGGACATCCAGCTACCATCATGGTGACCAAT  
CCACACTACCAGGGCCAAGTGAGCTTCCTGGACCCAGCTATTCCCTGCATATCAGCAATCTG  
AGCTGGGAGGATTCAGGGCTTTACCAAGCTCAAGTCAACCTGAGAACATCCAGATCTCTACC  
ATGCAGCAGTACAATCTATGTGTCTACCATCCTAACTATGCTTCTGAGAAGCCTTCAACAGCC  
TTCTGCCTCCTGGCCAAGGGATTGCTCATCTTCTTGCTCTTGTAATTCTGGCCATGGGACTC  
TGGGTCATCCGAGTCCAGAAAAGACACAAAATGCCAAGGATGAAGAACTCATGAGAAACAGA  
ATGAAATTGAGGAAGGAGGCAAAGCCTGGCTCCAGCCCTGCCTGACTGCTCCTTGGGAACCCC  
AGTCCTGAGCTTGGTTTCTTCCCAGCACCCAGAGAATCCTTCCTCAGCTCTCTTCTTTCCAGG  
GGAAGGAGGTGCTCAGGGGTGGGTATCCAGAGAGCCATACTTCTGAGGGAAGACTGGCTGGCA  
ATAAAGTCAAATTAAGTGACCACA

506/615

**FIGURE 502**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96879
><subunit 1 of 1, 198 aa, 1 stop
><MW: 22584, pI: 9.40, NX(S/T): 1
MCAFPWLLLLLLLLQEGSQRRLLWRWCGSEEVVAVLQESISLPLEIPPDEEVENIIWSSHKS
LATVVPGKEGHPATIMVTNPHYQGQVSFLDPSYSLHISNLSWEDSGLYQAQVNLRTSQIS
TMQQYNLCVYHPNYASEKPSTAFCLLAKGLLIFLLLVLAMGLWVIRVQKRHKMPRMKKL
MRNRMKLRKEAKPGSSPA
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-18

**Transmembrane domain:**

Amino acids 144-165

**N-glycosylation site:**

Amino acids 99-103

**N-myristoylation site:**

Amino acids 106-112

507/615

**FIGURE 503**

ACGGGCCCGCAGCGGCAGTGACGTAGGGTTGGCGCACGGATCCGTTGCGGCTGCAGCTCTGCAG  
TCGGGCCCGTTCCCTTCGCCGCCGCCAGGGGTAGCGGTGTAGCTGCGCAGCGTCGCGCGCGCTAC  
CGCACCCAGGTTTCGGCCCCGTAGGCGTCTGGCAGCCCCGGCGCCATCTTCATCGAGCGCC**ATGGC**  
CGCAGCCTGCGGGCCGGGAGCGGCCGGGTACTGCTTGCTCCTCGGCTTGCAATTTGTTTCTGCT  
GACCGCGGGCCCTGCCCTGGGCTGGAACGACCCTGACAGAATGTTGCTGCGGGATGTAAAAGC  
TCTTACCCTCCACTATGACCGCTATACCACCTCCCGCAGGCTGGATCCCATCCCACAGTTGAA  
ATGTGTTGGAGGCACAGCTGGTTGTGATTCTTATACCCCAAAGTCATACAGTGTGAGAACAA  
AGGCTGGGATGGGTATGATGTACAGTGGGAATGTAAGACGGACTTAGATATTGCATACAAATT  
TGGA AAAACTGTGGTGAGCTGTGAAGGCTATGAGTCCTCTGAAGACCAGTATGTACTAAGAGG  
TTCTTGTGGCTTGGAGTATAATTTAGATTATACAGAACTTGGCCTGCAGAACTGAAGGAGTC  
TGGAAGCAGCACGGCTTTGCCTCTTCTCTGATTATTATTATAAGTGGTCCTCGGCGGATTCT  
CTGTAACATGAGTGGATTGATTACCATCGTGGTACTCCTTGGGATCGCCTTTGTAGTCTATAA  
GCTGTTCTTGAGTGACGGGCAGTATTCTCCTCCACCGTACTCTGAGTATCCTCCATTTTCCCA  
CCGTTACCAGAGATTCACCAACTCAGCAGGACCTCCTCCCCAGGCTTTAAGTCTGAGTTCAC  
AGGACCACAGAATACTGGCCATGGTGCAACTTCTGGTTTTGGCAGTGCTTTTACAGGACAACA  
AGGATATGAAAATTCAGGACCAGGGTTCTGGACAGGCTTGGGAACTGGTGGAATACTAGGATA  
TTTGT TTGGCAGCAATAGAGCGGCAACACCCCTTCTCAGACTCGTGGTACTACCCGTCCTATCC  
TCCCTCCTACCCTGGCACGTGGAATAGGGCTTACTCACCCCTTCATGGAGGCTCGGGCAGCTA  
TTCGGTATGTTCAAACCTCAGACACGAAAACCAGAACTGCATCAGGATATGGTGGTACCAGGAG  
ACG**ATAA**AGTAGAAAGTTGGAGTCAAACACTGGATGCAGAAATTTTGGATTTTTCATCACTTT  
CTCTTTAGAAAAAAAGTACTACCTGTTAACAATTGGGAAAAGGGGATATTCAAAGTTCTGTG  
GTGTTATGTCCAGTGTAGCTTTTGTATTCTATTATTTGAGGCTAAAAGTTGATGTGTGACAA  
AATACTTATGTGTTGTATGTCAGTGTAAACATGCAGATGTATATTGCAGTTTTTGAAAGTGATC  
ATTACTGTGGAATGCTAAAAATACATTAATTTCTAAAACCTGTGATGCCCTAAGAAGCATTAA  
GAATGAAGGTGTTGTACTAATAGAACTAAGTACAGAAAATTTCAGTTTTAGGTGGTTGTAGC  
TGATGAGTTATTACCTCATAGAGACTATAATATTCTATTTGGTATTATATTATTTGATGTTTG  
CTGTTCTTCAAACATTTAAATCAAGCTTTGGACTAATTATGCTAATTTGTGAGTTCTGATCAC  
TTTTGAGCTCTGAAGCTTTGAATCATTCAGTGGTGGAGATGGCCTTCTGGTAACTGAATATTA  
CCTTCTGTAGGAAAAGGTGGAAAATAAGCATCTAGAAGGTTGTTGTGAATGACTCTGTGCTGG  
CAAAAATGCTTGAAACCTCTATATTTCTTTCGTTTCATAAGAGGTAAAGGTCAAATTTTCAAC  
AAAAGTCTTTTAATAACAAAAGCATGCAGTTCTCTGTGAAATCTCAAATATTGTTGTAATAGT  
CTGTTTCAATCTTAAAAAGAATCA

508/615

**FIGURE 504**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96889  
><subunit 1 of 1, 339 aa, 1 stop  
><MW: 36975, pI: 7.85, NX(S/T): 1  
MAAACGPGAAGYCLLLGLHLFLLTAGPALGWNDPDRMLLRDVKALTLHYDRYTTSRRLDPIPQ  
LKCVGGTAGCDSYTPKVIQCQNKGDGYDVQWECKTDLDIAYKFGKTVVSCEGYESSEDQYVL  
RGSCGLEYNLDYTELGLOKLKESGKQHGFASFSDYKKWSSADSCNMSGITIVVLLGIAFVV  
YKFLSDGQYSPPPYSEYPPFSHRYQRFTNSAGPPPPGFKSEFTGPQNTGHGATSGFGSAFTG  
QQGYENSGPGFWTGLGTGGILGYLFGSNRAATPFSDSWYYPSYPPSYPGTWNRAYSPHGGSG  
SYSVCSNSDTKTRTASGYGGTTRR

**Signal peptide:**

amino acids 1-30

**Transmembrane domain:**

amino acids 171-190

**N-glycosylation site.**

amino acids 172-176

**Glycosaminoglycan attachment sites.**

amino acids 244-248, 259-263, 331-335

**Tyrosine kinase phosphorylation site.**

amino acids 98-106

**N-myristoylation sites.**amino acids 68-74, 69-75, 131-137, 241-247, 247-253, 266-272,  
270-276, 278-284, 312-318

509/615

**FIGURE 505**

GCAAAAGGAAGGGAGGGAAGCACTCCATCATCTCACTGGGAAGAACGGCACGGGCATACCTGC  
AGCTACTGGGGTTCCACTGGGCTTGAGGGTCGATTTTTTCACCTTTTGAAGGACAAG**ATG**CATT  
GGAAGATGTTGCTGCTTCTGCTGTTGTATTACAATGCTGAGGCTTCTATGTGCCACAGGTGGA  
GCAGGGCTGTGCTCTTCCCTGCCGCCCACCGGCCAAAGAGGTCCTCATCACTGCCATTGAACC  
CAGTCTTGCAGACCTCCCTGGAGGAGGTGGAGCTGCTCTACGAGTTCTGCTGGCCGAACCTTG  
AGATCAGCCCTGACCTGCAGATCTCCATCAAGGACGAGGAGCTGGCCTCCTTGCGGAAGGCCT  
CAGACTTCCGCACCGTCTGCAACAACGTCATCCCCAAGAGCATCCCAGACATCCGCCGGCTCA  
GCGCCAGCCTCTCCAGCCACCCTGGCATCCTCAAGAAAGAAGACTTTGAAAGGACAGTGCTGA  
CCCTGGCCTACACAGCCTACCGCACAGCCCTGTCCACGGCCATCAGAAGGACATCTGGGCGC  
AGTCCCTCGTTAGCCTCTTCCAGGCCCTGAGGCACGACTTGATGCGCTCCTCACAGCCGGGAG  
TACCTCCCT**TGA**GAGACTGGCCACACCAGGACCTCAGAGCAGGGACCAGCACAGTAATCCAGA  
AAGTCTTCATTCTCTACTCCATTTACAGAGACCAGCAACAAAACACTTACCGCTGACACAGAG  
CAGCAGAGATCAAACAGTAACCCCGATGCTCTTTTCTCCTTGAGTTTCTGGAAGACACATC  
TGATTTCATGCCATCATGTGACCTGGGCTGGAAGAAAGGGCTGGAATGGTCATTCAAGACGCCT  
CCATGGGCAGAATGGTTTGCCTATGGCAGGCAGAATTCTGATATGCTTCAACCCAGAGCAGTG  
GCCACACACTCAAGAGTGAGAACAGGCGTGAGCCACCGTGCCTGGCCCAGGATCTAAAACTT  
TCTAAGTTTCTCCATCGTTGGCATCCTCACAGCTATCTCCAATGTCACTCAAGAGACATCAA  
CAGACATTTAACTGCTGCAGACTTCATTGCTCTGTACCTCACCTTGAATCTAACAAATCAA  
GTATTTCTGCAGGTCCAATGGTCTAAAATCAAATGCTTGTTAAATGACTTTTTTACAACACCCCTT  
ACTTTCCTAATCCATTTCAATCTTATTTTTTTTTTATTGTGGTAAAAAACACATCACGTAAATG  
TACCATCTTAACCATTTTTTAAGCATATGGTACAGCAGTGTTAACTCCATGCATGTTGTGAAAC  
AGACCCCCGGAACCTTCTCATCTTGTAATTCTGAAGTTCTATACCCACCGAACAACCTCCTCTT  
TTCCCCCTCCCCCTGCCTGCCCCAGCTCTTGGCACCATTATTCTGCTTTCTGTTTTTGGAGAGT  
CTGACTACTTAAGATACCTCATACAAGCGGGATCTGGCTTACATTTCTTGAGCATTGTATTCT  
GGAAAAGTGTTTCTTCTCTGAAAAATGGGTAGAGTTCTGAAGGAGAACTACTGGTCTTATT  
GTACACTTGCTGTACCTATTTTTTATTTAAACAAATATTCATCTATGGTATAATAAAGATGTCAT  
GGTTGGAAAAAAAAAAAAAAAAAAAAAAAAA



510/615

**FIGURE 506**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96893
><subunit 1 of 1, 173 aa, 1 stop
><MW: 19733, pI: 8.78, NX(S/T): 0
MHWKMLLLLLLLYYNAEASMCHRWSRAVLFPAAHRPKRSSSLPLNPVLQTSLEEVELLYEF
LLAELEISPDQLQISIKDEELASLRKASDFRTVCNNVIPKSIPDIRRLSASLSSHPGILKK
EDFERTVLTLAYTAYRTALSHGHQKDIWAQSLVSLFQALRHDLMRSSQPGVPP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-17

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 36-40;84-88;105-109

511/615

**FIGURE 507**

GGCGGCGGGCTGCGCGGAGCGGCGTCCCCTGCAGCCGCGGACCGAGGCAGCGGCGGCACCTGC  
CGGCCGAGCAATGCCAAGTGAGTACACCTATGTGAAACTGAGAAGTGATTGCTCGAGGCCTTC  
CCTGCAATGGTACACCCGAGCTCAAAGCAAG**AT**GAGAAGGCCAGCTTGTTATTAAGACAT  
CCTCAAATGTACATTGCTTGTGTTTGGAGTGTGGATCCTTTATATCCTCAAGTTAAATTATAC  
TACTGAAGAATGTGACATGAAAAAATGCATTATGTGGACCCTGACCATGTAAAGAGAGCTCA  
GAAATATGCTCAGCAAGTCTTGCAGAAGGAATGTGTCCTCAAGTTTGCCAAGACATCAATGGC  
GCTGTTATTTGAGCACAGGTATAGCGTGGACTTACTCCCTTTTGTGCAGAAGGCCCCCAAAGA  
CAGTGAAGCTGAGTCCAAGTACGATCCTCCTTTTGGGTTCCGGAAGTTCTCCAGTAAAGTCCA  
GACCCTCTTGGAACCTCTTGCCAGAGCACGACCTCCCTGAACACTTGAAAGCCAAGACCTGTCG  
GCGCTGTGTGGTTATTGGAAGCGGAGGAATACTGCACGGATTAGAAGTGGGCCACACCCTGAA  
CCAGTTCGATGTTGTGATAAGGTTAAACAGTGCACCAGTTGAGGGATATTGAGAACATGTTGG  
AAATAAACTACTATAAGGATGACTTATCCAGAGGGCGCACCACTGTCTGACCTTGAATATTAT  
TCCAATGACTTATTTGTTGCTGTTTTATTTAAGAGTGTTGATTTCAACTGGCTTCAAGCAATG  
GTAAAAAAGGAAACCCTGCCATTCTGGGTACGACTCTTCTTTTGGGAAGCAGGTGGCAGAAAAA  
ATCCCACTGCAGCCAAAACATTTTCAGGATTTTGAATCCAGTTATCATCAAAGAGACTGCCTTT  
GACATCCTTCAGTACTCAGAGCCTCAGTCAAGGTTCTGGGGCCGAGATAAGAACGTCCCCACA  
ATCGGTGTCATTGCCGTTGTCTTAGCCACACATCTGTGCGATGAAGTCAGTTTGGCGGGTTTT  
GGATATGACCTCAATCAACCCAGAACACCTTTGCACTACTTCGACAGTCAATGCATGGCTGCT  
ATGAACCTTCAGACCATGCATAATGTGACAACGGAAACCAAGTTCCTCTTAAAGCTGGTCAAA  
GAGGGAGTGGTGAAAGATCTCAGTGGAGGCATTGATCGTGAATTT**TGA**ACACAGAAAACCTCA  
GTTGAAAAATGCAACTCTAACTCTGAGAGCTGTTTTTGACAGCCTTCTTGATGTATTTCTCCAT  
CCTGCAGATACTTTGAAGTGCAGCTCATGTTTTTAACCTTTTAATTTAAAAACACAAAAAAAT  
TTTAGCTCTTCCCACTTTTTTTTTTCTATTTATTTGAGGTCAGTGTTTGTGTTTGCACACCAT  
TTTGTAATGAACTTAAGAATTGAATTGGAAGACTTCTCAAAGAGAATTGTATGTAACGAT  
GTTGTATTGATTTTTAAGAAAGTAATTTAATTTGTAAACTTCTGCTCGTTTACACTGCACAT  
TGAATACAGGTAACATAATTGGAAGGAGAGGGGAGGTCACCTTTTGATGGTGGCCCTGAACCT  
CATTCCTGGTCCCTGCTGCGCTGCTTGGTGTGACCCACGGAGGATCCACTCCCAGGATGACGT  
GCTCCGTAGCTCTGCTGCTGATACTGGGTCTGCGATGCAGCGGCGTGAGGCCTGGGCTGGTTG  
GAGAAGGTCACAACCCCTTCTCTGTTGGTCTGCCTTCTGCTGAAAGACTCGAGAACCAACCAGG  
GAAGCTGTCCTGGAGGTCCCTGGTGGAGAGGGACATAGAATCTGTGACCTCTGACAACTGTG  
AAGCCACCCTGGGCTACAGAAACCACAGTCTTCCAGCAATTATTACAATTCTTGAATTCCTT  
GGGATTTTTTACTGCCCTTTCAAAGCACTTAAGTGTAGATCTAACGTGTTCCAGTGTCTGT  
CTGAGGTGACTTAAAAAATCAGAACAAAACCTTCTATTATCCAGAGTCATGGGAGAGTACACCC  
TTTCCAGGAATAATGTTTTGGGAAACACTGAAATGAAATCTTCCAGTATTATAAATTGTGTA  
TTTAA

512/615

**FIGURE 508**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA96897
><subunit 1 of 1, 362 aa, 1 stop
><MW: 41736, pI: 8.80, NX(S/T): 3
MRRPSLLLLKDILKCTLLVFGVWILYILKLNYTTEECMDKMKMHYVDPDHVKRAQKYAQQVL
QKECRPKFAKTSMAALLFEHRYSDLLPFVQKAPKDSEAESKYDPPFGFRKFSSKVQTLLE
LLPEHDLPEHLKAKTCRRCVVIGSGGILHGLELGHNTLNQFDVVIRLNSAPVEGYSEHVGN
KTTIRMTYPEGAPLSDLEYYSNDLFVAVLFKSVDFNWLQAMVKKETLPFWVRLFFWKQVA
EKIPLQPKHFRILNPVVIKETAFDILQYSEPQSRFWGRDKNVPTIGVIAVVLATHLCDEV
SLAGFGYDLNQPRTPPLHYFDSQCMAAMNFQTMHNVTTETKFLKLKLVKEGVVKDLSSGGIDR
EF
```

**Important features of the protein:****Transmembrane domains:**

Amino acids 11-27;281-297

**N-glycosylation sites:**

Amino acids 30-34;180-184;334-338

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 2-6;109-113;223-227

**N-myristoylation sites:**

Amino acids 146-152;150-156;179-185;191-197

513/615

**FIGURE 509**

GGGCGGACGCAGTGCAGTAAGAGCAGATGGGCGGACCCAAATTTCTTCGGCTTCACGATTTTG  
CCGAGGTCTAGCCCTGCATCCAGCCTTGAAACAGGGTGGGGAGGAGGCAGAAAGGGGAGGGAC  
TGCACTCCCTCTGAGCGTGCTAGCTCCGACTGCCTGACGGATCACCCCTCCGCTCCAACATGG  
CTAGTTCCCTCAACGCCGTGACTCAAGCCTGTTGTGCCAGGCAGGGCGCACTCAGCAGCGCAGC  
CCCACAGGTGGCGAAGGCTCCGCGAGAGGGTTCCCGCCAGGCTAGACAGTGGAGTGCCGCACA  
GCGCGCCTTCCAGCCTCGCAGCCGCCACCCTAGCGGTTCCGACCCGGCGCCAGCAGGCCTGCT  
TGGTCGATCTTCGAGCCAAAG**ATG**CGGCGAGGCTGGAAGATGGCTCTGTCTGGGGGGCTGCGG  
TGCTGCCGCCGGGTACTGTCCTGGGTGCCAGTGCTCGTTATTGTCCTCGTCGTGCTCTGGTCC  
TACTATGCCTACGTCTTTGAACTCTGCCTGGTTATTTACCTCATACTCTACCATGCCATCTTT  
GTGTTCTTTACCTGGACCTACTGGAAGTCTATCTTTACACTCCACAGCAGCCAAACCAGAAG  
TTCCACTTGTCCTACACAGACAAGGAGCGCTATGAAAATGAAGAAAGACCTGAGGTCCAGAAG  
CAGATGCTTGTTGATATGGCCAAAAAGCTACCGGTTTACACAAGAACTGGAAGTGGAGGTCAG  
TTCATCCAAAGGCAGCTAGAGAGGCAGCTCAGCAAGTATCTCAGAAAGGCTAAGTCATATATG  
TTCTCAAAC**TAG**CCCTTTTTTTTCCCTCCCATCTTCTGAAAACCACTATGGAGATTTTTCTCCA  
CATTTTATTTCTAAAAAATTTTAAACACATATCAAAGCTGGAAGAATTGTATAGTAAACAAAC  
TGTATACCCCAAACCTGGATTCTTCTGCTAACATTTTTCTGTGTTGCTATATCACATATCTATC  
CACATATGCATACCTCTATTTATCTTTCGTCAAGCCATCTTATGTTTCTGATGCATTTCAAAG  
TAAATAGCTGACATCAGTAAGACATCTACCTAAATATTTTATTCTGTTTTGTAAAATTTACA  
TACAAAAACATGCATAATCTTAAGGGTACCATTCCATGTATTTTGAAAAGTGTACACATCTGT  
GTAACATAACCCCAATAAAATTGCCATCACCTCAG

514/615

**FIGURE 510**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA98564
><subunit 1 of 1, 143 aa, 1 stop
><MW: 17255, pI: 9.99, NX(S/T): 0
MRRGWKMALSGGLRCCRRVLSWVPVLVIVLVVLWSYYAYVFELCLVIYLLILYHAIFVFFT
WTYWKSIFTLPPQPNQKFHLSYTDKERYENEERPEVQKQMLVDMAKKLPVYTRTGSGGQF
IQRQLERQLSKYLRKAKSYMFSN
```

**Important features of the protein:****Transmembrane domain:**

Amino acids 24-45

**N-myristoylation sites:**

Amino acids 11-17;12-18

515/615

**FIGURE 511**

CAGCCGGGGCG**ATG**GCGGGGCTCTGGCTGGGGCTCGTGTGGCAGAAGCTGCTGCTGTGGGGCG  
CGGCGAGTGCCCTTTCCCTGGCCGGCGCCAGTCTGGTCCTGAGCCTGCTGCAGAGGGTGGCGA  
GCTACGCGCGGAAATGGCAGCAGATGCGGGCCATCCCCACGGTGGCCCGCGCCTACCCACTGGTG  
GGCCACGCGCTGCTGATGAAGCCGGACGGGCGAGGTAAGGGCCGGCGCTCCTCCTGGAGCGCA  
ACGGGGTCCGCAGCCCCGTTCACCCCTCCGATCAGCCAGGAACCCGCTGCTTGTGGCGCTGG  
CCGCAGGAGAGAGGAGCCTGTCACCCTGTGGAGAATGCACTCCCAGTTCTAGTCGTTGCCCT  
TGGCACCCGCCGACACTGCTAGTGCCCCATCCCAAAGTGAGCATTCTTTGTGTGTAGCACA  
GGATGCGGTATTTCCAAACCCCTGCCCTCGGTCTTTTCCACCTCACCGCTGCTCAGCTCTCA  
AAGCCCTGCCGTTTCTCCTGCCTTGGCTTGGGAAGCCT**TAG**GAACAGAAGCTCCCTGGGAGC  
ACAGAGCGGTTTTAAACTGGCCAACACCTTAACGCCAGAGCCGCCCTCCTCTCGCTGCCACT  
TTGGAAAATAAGAGACTAGAGATTCAGTGGACGCTTCCTCCCGGCATCACAAGACTTGACTGC  
TGCTTCAGTTCCCGCTTGACCTTCATACTTTAGCCCTTTAAAGGATGTTACATAATAACAATT  
AAGAGACGGCAGGGCCCTTCAGGCAGACTTCTTTGGAGGGTGTCAAACGCCCTTGTTTATTAAAG  
AGTGAATTTTTTAATTAAAATCATGTTTTAAACAGAGATGGACATTTTATTGATGGAAAAA  
ATCACGTTAAGTTAGAAAGCTCTCAAAGTACCTGGTATTTACAACCTCCCTGTCAGGGAGGGC  
GAACTCGATCTCAGAGTTTTATTTTCATCAGGGATTACGTTGAGGTACCCAGAAATGAGAAGA  
TTTGCCCAAAATGGCATATTTTAAATTTGGCCAGACCAGAACCCAGTTTCCTCTGGGATTAT  
TTGTTAGTAATCGTTTTACAGGCTGAGCATTAACTAACTCCAAAGCTTGAAGGACTTTTTCTC  
ATTTTCACTTGTTTTCTCTAATAAAAATAATGCTGTAATTTCACTTCACAAGATGAGGCCTC  
ATGGAAGAGTGTTTTACCAAATATTTAAAATACTTTGACAGAAAAAATCAAGCGAACTCTTT  
GCCAACCAAATATCATCATGACTGATGTAACAAGTAATCCAACACAGATATGAAAATCACTGG  
TAAAATCATCTCAGTTAATTCTAAAAGCAGAGCTAACCACCCCTTTTGTCTTAAGGCTTTAT  
GGTATTAATAAAATAAACTGTACAAAATATAGATTTTCCCCTATCCCCTACCCCTGGAAAGTA  
ATATACTGAAGTCTCATCATACTGTTTTGGGGATTCCAGTAATTAATCTCTAGTGAACAAA  
GACCTGTTTCAAAACAACCTGTGAGCTGACTGGACTATTTAAAGTAATTCTCCTTGTAGTCAC  
TTTCAGAGTGAAGACAATGACGAATACTGTCTTTTACAAAGGGACTTTTTATTCCACCAACAA  
ATTCTGGATTTTGGCATCAGGAAAACCACTGTTCAATTTCCAACACTATATCCAAGTTGTTTG  
AGAAATTATTTAAACTCTTTAACTTAGAGGGTTTTCTTTCTCCTTTACTTGTTAAAGTGACT  
ATATTACAGAGTCACTTTAAGGATTAAATTTATTGCATGCAAAGTTTCTAGATCACTGTCTAG  
AAGTCAGTTAGAGTAAGTTCTTTAGTTGTCAATCAAGCATTTAGTAAGGCCCTGCTTTGTGCC  
CAGTGTTGACCTCAACAAAGTTGGGGATATCAGAATATTCTAAGATACAGTCGTTGTTCCCAA  
GAATCTTGCTTTTACATACAAGAGGTGTTGCGTTTCATTTTGCGGCTAATGTCCAAACGCTG  
GCCTCAGCCATTTACCTTGAAGATTGCAGTTGGCTTCCAACCTGGCCTCTAACTCTAATCTA  
GCATTTTCCAGTCCATTGTGACAAAGTCTGCCTTCCCAGCTACTCCCAGCTGTTGGACCTGC  
TGCCTTAGAACCACAGATTGGTACCTCGTGCC

516/615

**FIGURE 512**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA107443
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19353, pI: 10.97, NX(S/T): 0
MAGLWLGVLVWQKLLLLWGAASALSLAGASIVLSLLQRVASYARKWQQMRPIPTVARAYPLV
GHALLMKPDGRGKGRSSWSATGSAAPFPPSDQPGTRCLWRWPQERGACHPVENALPVLV
VAPWHPPTLLVPHPKVSIFFVCSTGCGISKPLPSVFSHLTAAQLSKPCRFLLPWLGKP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-25

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 75-79

**N-myristoylation sites:**

Amino acids 3-9;17-23;145-151

**Amidation site:**

Amino acids 73-77

**Leucine zipper pattern:**

Amino acids 8-30

517/615

**FIGURE 513**

GGCGGCTGGACGAGGACGCTCAGAGCCCAGCTCTCGAGAGTTCAAGCAACCGACGGTTCC  
CCACTGCTCCCAGGAGCGGTTACCTGGGCACCTCTGTGCCCCCTCCTTCCTGTTGGGGCCCA  
GGCCGAGGACCTGCCAGTAGGGCTCAGTTGCCTGGAGCCCGTTAGCCCATCCCCAGTT  
CACTTTGCTTGTGGGATCTCCCCGTTGCTCCTGCCCCTGGACTGAGTGGCAGGCCATCCT  
ACAAGCACCCGGACACTTGACATCAGTGGTGTCAAGACAACCTCTAAGAAGGTTTTCCGTG  
ATCCTGCAAGCCCTGCCTTCCTTCCTGGGATCCTGCCTTCAATTTGATTGCACAGGTACC  
ACAGCAAGCCAGTGCTGTGTGCTCCGAGTTCCAGGGCGTCCTCCAGCTCAGCCACTGCAC  
TGAGAACAATGGACTCTCTGTGGGGCCAGGAGCCGGGAGTCACCCCTTTGGGGTCCACAA  
CACCCGGCTGTCCCCAGACTTGTGTCCAGGGAAGATAGTGTGAGGGCCCTCAAGGAGAG  
CGGGGCAGGGATGCCTGAGCAGGACAAGGACCCTAGAGTCCAAGAGAATCCTGGTGATCA  
GAGAAGGGTCCCCGAGGTACCCGGGGATGCACGGTCTGCATTTGCGCCCCCTGCGGGACAA  
TGGAGGCCCTCTCTCCCTTTGTGCCCGGGCCCCGGGCCCTCTGCAGACAGACCTCCATGCCCA  
GAGGTGAGAAATCAGATATAACCAGACATCCCAGACCTCCTGGACGAGCTCCTGCACCAA  
CCGAAATGCCATCTCCAGCTCCTACAGCTCCACGGGAGGCTTGCTGGGGCTAAAGCGGAG  
GAGGGGGCCAGCCTCATCCCACTGCCAGCTGACCCTCAGTTCTCAAAGACAGTGAGTGA  
GGACAGGCCTCAGGCTGTCTCTCAGGTACACCCAGTGTGAAAAGGCAGCAGATATAGC  
ACCAGGGCAGACACTCACCTCAGGAATGACTCCTCCACATCCGAGGCCTCTAGGCCCAG  
TACACACAAGTTTCCCCTGCTGCCACGCAGGCGAGGGGAGCCTTTGATGCTGCCACCTCC  
CTTAGAGCTGGGGTACCGGGTCACTGTTGAAGACCTGGACCGGGAGAAGGAGGCGGCCTT  
CCAGCGCATCAACAGTGCAGTGCAGGTTGAGGACAAGGCCATCTCGGACTGCAGACCCCTC  
ACGGCCTTCCCACACTTTGTCTCTCACTTGCAACAGGGGCTTCTGGTCTGCCTGCCGTTTC  
TAAAGCACCCAGTATGGATGCACAGCAGGAGACACACAAGTCCCAAGACTGCCTGGGCCT  
ACTGGACCCCTTAGCATCTGCTGCAGGGGTCCCCTCTACAGCTCCCATGTCTGGGAAGAA  
GCACAGACCACCAGGCCCCCTGTTCTCCTCCTCAGATCCCCTTCTGCCACCTCTTCTGA  
TTCCCAGGACTCAGCCCAGGTACCTCGCTGATTCTCTGCCCCCTTCCCAGCTGCAAGCAT  
GGATGCGGGCATGAGAAGAACAAGGCATGGCACTTCTGCTCCTGCAGCTGCCGCAGCAGC  
CCCTCCCCGCTCCACATTGAACCCACGTTGGGGTCACTACTGGAGTGGATGGAGGCCCT  
TCACATTTCTGGGCCTCAGCCACAGCTGCAGCAGGTGCCAGAGGTGAGAACCAGAGATC  
CCAGACCTCCTGGACCAGCTCGTGCCCCAAATGAATGCCATCTCGAGCCCCCTACAGCTC  
TACGGGAGGCCTCCCGGAACAAAAGCGGAAGAGGGGCCAGCCTCATCCCACTGCCAGCTG  
ACCCTCAGTTCTCAAACACAGTGAGTGAGGACGGACCTCAGGCTGTCTCTCGGGTCA  
ACCCAGTGTGAAAAGACGGCAGATACAGCACCAGGGCAGACACTCGCCTCCAGGGGTGGC  
TCCCCCAGATCCCAGGCCTCTAGGCCCCGTATATGCAAGTTTCCCCTGCTGCCACGCAGG  
CGAGGGGAGCCTTTGATGCTGCCACCTCCCTTAGAGATGGGGTACCGGGTCACTGCTGAA  
GACCTGGACCGGGAGAAGGAGGAGGCATTCCAGCGCATCAACAGTGCAGTGCAGGTTGAG  
GACCAGGCCATCTAGGACTGCAGACCCTCACGGCCTTCCCACACTTTGTCTCACTTGCA  
ACAGGGGCTTCTGGTCTGCCTGCCGTTTCTAAAGCACCCAGTATGGATGCACAGCAGGAG  
ACACACAAGTCCCAAGACTGCCTGGGCCTAGTGGCCCCCTGCATCTGCTGCACAGGCCT  
GTAGTCCCAGCTACTTGGGAGGCTGAGGCAGGAGAACGGCATAAACCGGGGAGGCAGAGC  
TTGCAGTGAGCTGAGATCGCGCCACTGCACTCCAGCCTGGGTGACAGAGCGAGACTCCGT  
CTC



518/615

**FIGURE 514**

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA107786
><subunit 1 of 1, 428 aa, 1 stop
><MW: 45450, pI: 9.28, NX(S/T): 3
MDSLWGPAGSHPPFGVHNTRLSPDLCPGKIVLRALKESGAGMPEQDKDPRVQENPGDQRR
VPEVTGDARSAFRPLRDNGGLSPFVPGPGPLQTDLHAQRSEIRYNQTSQTSWTSSCTNRN
AISSSYSSSTGGLLGLKRRRGPASSHCQLTLSSSKTVSEDRPQAVSSGHTQCEKAADIAPG
QTLTLRNDSSSTSEASRPSTHKFPLPRRRGEPLMLPPPPELGYRVTVEDLDREKEAAFQR
INSALQVEDKAISDCRPSRPSHTLSSLATGASGLPAVSKAPSMDAQQETHKSQDCLGLLD
PLASAAGVPSTAPMSGKKHRPPGPLFSSSDPLPATSSDSQDSAQVTS LI PAF FPAASMDA
GMRRTRHGTSAPAAAAAAPPRSTLNPTLGSLLLEWMEALHISGPQPQLQOVPRGQNQRSQT
SWTSSCPK
```

**Important features of the protein:****N-glycosylation sites:**

Amino acids 105-109;187-191

**Glycosaminoglycan attachment site:**

Amino acids 38-42

**N-myristoylation sites:**

Amino acids 15-21;130-136;180-186;307-313;361-367

**Amidation site:**

Amino acids 315-319

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 106-117

519/615

**FIGURE 515**

GTCAGGGCCAGGGTGAGCGCCCGACTCCGAGCTGTCCCCGCTCCCGGCGCGGCGCTCCGCTCT  
CAGCCACCTCACGGCTGCCAGGAGTGCGCGGGAGTTTGCCCCGGAGCGCGGGGAAGTTTCCTC  
CGAAGCTGCGCTCCTGGAACAGCAGCACCTGCAAGCGCCCGGCAGCGGCCCGCGAGGTTACTT  
TATGGAATTGGGCTCTTAGAGAACAAGAAAAGACTGAAGTTTACGGGAAAACAAATCATGTG  
GTCTTCAGATTCTGAAATAAGGAGAA**ATG**CAGCCATCTGAAATGGTCATGAACCCCAAACAAG  
TCTTCCTCTCTGTGCTGATATTTGGAGTAGCTGGGCTACTCCTCTTCATGTATTTGCAAGTCTGG  
ATTGAAGAACAACATACAGGGAGAGTGGAGAAGAGAAGAGAACAAAAAGTAAGTTCAGGATGG  
GGACCAGTGAAGTACTTGCGGCCTGTACCCAGAATCATGAGTACAGAAAAAATCCAGGAACAT  
ATCACCACCAGAACCCCAAGTTTACATGCCTGAGGATGTACGAGAAAAAAGGAAAATCTT  
CTACTCAATTCTGAGAGATCTACTAGGCTCTTAACAAAGACCAGTCATTACAAAGGAGGGGAT  
CAAGCTTTAAGTAAGTCCACAGGGTACCAACAGAGAAGTTGATTGAAAAACGTCAAGGAGCT  
AAGACTGTTTTTAAACAAGTTCAGCAACATGAATTGGCCAGTGGACATTCACCCTTTAAACAAA  
AGTTTAGTCAAAGATAATAAATGGAAGAAAAGTGAAGAGACCCAAGAGAAACGAAGTCTTTT  
CTTCAGGAGTTTTTGCAAGAAATACGGTGGGGTGAGTCATCATCAGTCACATCTTTTTCATACA  
GTATCCAGAATCTATGTAGAAGATAAACACAAAAATCTTATATTGTGAGGTACCTAAGGCTGGC  
TGTTCCAATTGGAAAAGAATTCTGATGGTACTAAATGGATTGGCTTCTCTGCATACAACATC  
TCCCACAATGCTGTCCACTACGGGAAGCATTGGAAGAAGCTAGATAGCTTTGACCTAAAAGGG  
ATATATACCCGCTTAAATACTTACACCAAAGCTGTGTTTGTTCGTGATCCCATGGAAAGATTA  
GTATCAGCCTTTAGGGACAAATTTGAACACCCCAATAGTTATTACCATCCAGTATTCGGAAAG  
GCAATTATCAAGAAATATCGACCAAATGCCTGTGAAGAAGCATTAATTAATGGATCTGGAGTC  
AAGTTCAAAGAGTTTATCCACTACTTGCTGGATTCCCACCGTCCAGTAGGAATGGACATTCAC  
TGGGAAAAGGTCAGCAAACCTCTGCTATCCGTGTTTGATCAACTATGATTTTGTAGGGAAATTT  
GAGACTTTGGAAGAAGATGCCAATTACTTTTTACAGATGATCGGTGCTCCAAAGGAGCTGAAA  
TTTCCCAACTTTAAGGATAGGCACCTCTCCGATGAAAGAACCAATGCTCAAGTCGTGAGACAG  
TATTTAAAGGATCTGACTAGAACTGAGAGACAATTAATCTATGACTTTTATTACTTGACTAT  
TTAATGTTTAATTATACAACCTCACTTTT**GTAG**TTTGCATTCATTTTCTAAAACCTGTATAT  
ACTTAATGATGATAAGTTCAAATCAGCTGTAATTTTTCTATAATTCTCTGTATGACAGAAATT  
TAACCAAGTGCAGTTGTCTTGATTTAATGTAGATTTTTACCAAATAGTATGACACCAATTGGC  
ACAAAGTTATAGGAAAATCACCTACAGGAGATGTAACAACCTTGAGTTGCTCTAAAATGTTTG  
GAAAAGAGCTGCTTTTGCAATTATGAATTATATTGTTGAAGCAATAACCTAGCCAGCTGTTGCA  
TTAGCTAAAGCAGCCTCTTGCAATGGTAGGAAAAAAGGATCTCAAATAGCATGAGTGTATGTC  
TATATCCTGAAATTTATTGTCTAAAATGCATGAATATATTTTTAGCAGTCTGTGGCATATTAA  
TCAAACGTGTTGAATTGTTTTCTTACACCCTGGAAATCTTCTATCAACTATAATGATAAATCC  
ATTTTGAAGTGATATTTTGGACTTAGGCATTTTACTTTAGATTGGAAGGCATTATGTGATTTACA  
ATATGAGAATATAGCAGAAAAACCA

520/615

**FIGURE 516**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108682
><subunit 1 of 1, 443 aa, 1 stop
><MW: 52021, pI: 9.63, NX(S/T): 4
MQPSEMVMNPKQVFLSVLIFGVAGLLLFMYLQVWIEEQHTGRVEKRREQKVTSGWGPVKY
LRPVPRIMSTEKIQEHITNQNPKFHMPEDVREKKENLLLNSERSTRLLTKTSHSQGGDQA
LSKSTGSPTEKLIKRGAKTVFNKFSNMNWPVDIHPLNKS LVKDNKWKKT EETQEKRRS
FLQEFCKKYGGVSHHQSHLFHTVSRIYVEDKHKILYCEVPKAGCSNWK RILMVLNGLASS
AYNISHNAVHYGKHLKKLDSFDLKG IYTRLNTYTKAVFVRDPMERLVSAFRDKFEHPNSY
YHPVFGKAI IKKYRPNACEEALINGSGVKFKEFIHYLLDSHRPVGMDIHWEK VSKLCYPC
LINYDFVGKFETLEEDANYFLQMIGAPKELKFPNFKDRHSSDERTNAQVVRQYLKDLTRT
ERQLIYDFYYLDYLMFNYYTTPLL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-24

**N-glycosylation sites:**

Amino acids 159-163;243-247;324-328;437-441

**Glycosaminoglycan attachment site:**

Amino acids 53-57

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 177-181

**Tyrosine kinase phosphorylation site:**

Amino acids 329-337

**N-myristoylation sites:**

Amino acids 116-122;236-242

521/615

**FIGURE 517**

GGAACCTCCAGGCACCCTGTGTGGCCGCACTGCTCCCTCTGGCCCAACCATGCCTCTGTCCAGCCACCTGCTGC  
CCGCTTGGTCTGTTCCTGGCAGGGTCTCAGGCTGGGCTGGGTCCCAACCACTGCAGGAGCCCTGGCCAGG  
CCGTGTGCAACTTCGTGTGTGACTGCAGGGACTGCTCAGATGAGGCCAGTGTGGTTACCACGGGGCCTCGCCCA  
CCTGGGCGCCCCCTTCGCTGTGACTTCGAGCAGGACCCCTGCGGCTGGCGGGACATTAGTACCTCAGGCTACA  
GCTGGCTCCGAGACAGGGCAGGGGCCGCACTGGAGGGTCTTGGGCTCACTCAGACCACACACTGGGCACCGACT  
TGGGCTGGTACATGGCCGTGGAAACCACCGAGGGAAAGAGGCATCCACCGCAGCCCTGCGCTCGCCAACCCTGC  
GAGAGGCAGCCTCCTCTTGAAGCTGAGGCTCTGGTACCACGCGGCTCTGGAGATGTGGCTGAACTGCGGGTGG  
AGCTGACCCATGGCGCAGAGACCTGACCCTGTGGCAGAGCACAGGGCCCTGGGGCCCTGGCTGGCAGGAGTTGG  
CAGTGACCACAGGCCGCATCCGGGGTGACTTCGAGTGACCTTCTCTGCCACCCGAAATGCCACCCACAGGGGCG  
CTGTGGCTCTAGATGACCTAGAGTTCTGGGACTGTGGTCTGCCACCCCCAGGCCAACTGTCCCCGGGACACC  
ACCAGTGCCAGAACAAGGTCTGCGTGGAGCCCCAGCAGCTGTGCGACGGGGAAGACAACCTGCGGGGACCTGTCTG  
ATGAGAACCCTACTCACCTGTGGCCGCCACATAGCCACCGACTTTGAGACAGGCCTGGGCCCATGGAACCGCTCGG  
AAGGCTGGTCCCGGAACCACCGTGCTGGTGGTCTGAGCGCCCTCTGGCCACGCGCTGACCACAGCCGGAACA  
GTGCACAGGGCTCCTTCCCTGGTCTCCGTGGCCGAGCCTGGCACCCCTGCTATACTCTCCAGCCCCGAATTCCAAG  
CCTCAGGCACCTCCAAGTGTCTGGTCTTCTATCAGTACCTGAGTGGGTCTGAGGCTGGCTGCCTCCAGCTGT  
TCTGCGAGACTCTGGGGCCCGGCCCCCCGGGCCCCCGTCTGCTGCGGAGGCGCGAGGGGAGCTGGGGACCG  
CCTGGGTCCGAGACCGTGTGACATCCAGAGCGCTACCCCTTCCAGATCCTCCTGGCCGGGCAGACAGGCCCGG  
GGGGCGTCTGGGTCTGGACGACCTCATCTGTCTGACCACTGCAGACCAGTCTCGGAGGTGTCCACCCTGCAGC  
CGCTGCCTCCTGGGCCCCGGGCCCCAGCCCCCAGCCCTGCCGCCAGCTCGCGGCTCCAGGATTCTTGAAGC  
AGGGGCATCTTGCCTGCGGGGACCTGTGTGTGCCCCGGAACAACCTGTGTGACTTCGAGGAGCAGTGCAGAGGG  
GCGAGGACGAGCAGGCCTGTGGCACCACAGACTTTGAGTCCCCGAGGCTGGGGGCTGGGAGGACGCCAGCGTGG  
GGCGGCTGCAGTGGCGCGTGTCTCAGCCCAGGAGAGCCAGGGGTCCAGTGCAGCTGCTGCTGGGCACTTCTGT  
CTCTGCAGCGGGCCTGGGGCAGCTAGGCGCTGAGGCCGGGTCTCACACCCCTCCTTGGCCCTTCTGGCCCCA  
GCTGTGAACTCCACCTGGCTTATTATTTACAGAGCCAGCCCCGAGAGGTCTCCTGTAACCTTTGAGCGGGACACAT  
GCAGCTGGTACCCAGGCCACCTCTCAGACACACACTGGCGTGGGTGGAGAGCCGCGGCCCTGACCACGACCACA  
CCACAGGCCAAGGCCACTTTGTGCTCCTGGACCCACAGACCCCTGGCCTGGGGCCACAGTGCACACCTGCTCT  
CCAGGCCCCAGGTGCCAGCAGCACCACGGAGTGTCTCAGCTTCTGGTACCACCTCCATGGGCCCCAGATTGGGA  
CTCTGCGCCTAGCCATGAGACGGGAAGGGGAGGAGACACACCTGTGGTGCAGGTGAGGACCCAGGGCAACCGCT  
GGCAGGAGCCCTGGGCCACCCCTTTCCACAGCCCTGGCTCCATGCCAGTACCAGTGTCTGTCGAGGGCCCTCC  
GGGACGGATACACGGCACCATGGCGCTGGACGATGTGGCCGTGCGGCCGGGCCCTGCTGGGCCCCAATTACT  
GCTCCTTTGAGGACTCAGACTGCGGCTTCTCCCTGGAGGCCAAGGTCTCTGGAGGCGGAGGCCAATGCCTCGG  
GCCATGCTGCCTGGGGCCCCCAACAGACCATACCACTGAGACAGCCCAAGGGCACTACATGGTGGTGGACAAA  
GCCAGACGCACTACCCCGGGGCCAGACGGCCTCCCTGACCTCCAAGGAGCACAGGCCCTGGCCAGCCTGCTT  
GTCTGACCTTCTGGTACCACGGGAGCCTCCGCAGCCAGGCACCCTGCGGGTCTACCTGGAGGAGCGCGGAGGC  
ACCAGGTGCTCAGCCTCAGTGCCACGGCGGGCTTGCTTGGCGCCTGGGCAGCATGGACGTGCAGGCCGAGCGAG  
CCTGGAGGGTGGTGTGTTGAGGCAGTGGCCGAGGCGTGGCACACTCCTACGTGGCTCTGGATGATCTGCTCCTCC  
AGGACGGGCCCTGCCCTCAGCCAGGTTCTGTGATTTTGAAGTCTGGCCTGTGTGGCTGGAGCCACCTGGCCGGG  
CCGGCCTGGGCGGATACAGCTGGGACTGGGGCGGGGAGCCACCCCTCTCGTTACCCCAAGCCCTGTGGACC  
ACACCTGGGCACAGAGGCAGGCCACTTTGCCTTCTTTGAACTGGCGTGTGGGCCCCGGGGGCCGGCCGCT  
GGCTGCGCAGCGAGCCTCTGCCGGCCACCCAGCCTCCTGCCTCCGCTTCTGGTACCACATGGGTTTCTCTGAGC  
ACTTCTACAAGGGGGAGCTGAAGGTACTGCTGCACAGTGTCTAGGGCCAGCTGGCTGTGTGGGGCGCAGGCGGGC  
ATCGGCGGCACCAAGTGGCTGGAGGCCAGGTGGAGGTAGCCAGTGCCAAGGAGTTCCAGATCGTGTGTAAGCCA  
CTCTGGGCGGCCAGCCAGCCCTGGGGCCATTGCCCTGGATGACGTGGAGTATCTGGCTGGGCAGCATTGCCAGC  
AGCCTGCCCCCAGCCCGGGGAACACAGCCGCACCCGGGTCTGTGCCAGCTGTGGTTGGCAGTGCCTCCTATTGC  
TCATGCTCCTGGTGTCTGGGACTTGGGGACGGCGTGGCTGCAGAAGAAGGGGAGCTGCCCTTCCAGAGCA  
ACACAGAGGCCACAGCCCTGGCTTTGACAACATCCTTTCAATGCGGATGGTGTACCCCTCCCGGCATCTGTCA  
CCAGTGATCCGTAGACCACCCAGACAAGGCCCCGCTTCTCAGTGACATCCAGCACTTGGTCAGACCCCTAGCC  
AGGACCGGACACCTGCCCCGCCCCAGGCTGGGACAGGCTCAGGATATGCTGAGGCTGGCGCTGGCCCTTCCC  
TGCCCTGTGCTGACTCTGTGCTCTGTGAATAAACACCCCTGGCCCATGAGGGCCGCCCAAAAAAAAAAAAAA  
AAAAAAAAAAAAAA

522/615

**FIGURE 518**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108684
><subunit 1 of 1, 1137 aa, 1 stop
><MW: 122776, pI: 6.00, NX(S/T): 4
MPLSSHLLPALVLFFLAGSSGWAVPNHCRSPGQAVCNFVCDRCDCDEAQCGYHGASPTL
GAPFACDFEQDPCGWRDISTSGYSWLRDRAGAALGPGPHSDHTLGTDLGWYMAVGTHRG
KEASTAALRSPTLREAASSCKLRLWYHAASGDVAELRVELTHGAETITLWQSTGPWGPWG
QELAVTTGRIRGDFRVTFSTRNATHRGAVALDDLEFWDCGLPTPQANCPPGHHHCQNKV
CVEPQQLCDGEDNCGDLSDENPLTCGRHIATDFETGLGPWNRSEGSRNHRAGGPERPSW
PRRDHSRNSAQGSFLVSVAEPTGPAILSSPEFQASGTSNCSLVFYQYLSGSEAGCLQLFL
QTLGPGAPRAPVLLRRRRRGELGTAWVRDRVDIQSAYPFQILLAGQTGPGGVVGLDDLILS
DHCRPVSEVSTLQPLPPGPRAPAPQPLPPSSRLQDSCKQGHLAGDLCVPPEQLCDFEEQ
CAGGEDEQACGTTDFESPEAGGWEDASVGRQLQWRRVSAQESQGSSAAAAGHFLSLQRAWG
QLGAEARVLTPLLGPSPGSCELHLAYYLQSQPREVSCNFERDTCWYPGHLSDTHWRWVE
SRGPDHDHTTGQGHFVLLDPTDPLAWGHSALHLSRPQVPAAPTECLSFYHLHGPQIGTL
RLAMRREGEETHLWSRSGTQGNRWHEAWATLSHQPGSHAQYQLLFEGLRDGYHGTMALDD
VAVRPGPCWAPNYCSFEDSDCGFSPGGQGLWRRQANASGHAAGPPTDHTTETAQGHYMV
VDTSPDALPRGQTASLTSKEHRPLAQPACLTFWYHGSLRSPGTLRVYLEERGRHQVLSLS
AHGGLAWRLGSMVDVQAERAWRVFEAVAAGVAHSYVALDDLLQLDGPQPGSCDFESGL
CGWSHLAGPGLGGYSWDWGGGATPSRYQPPVDHTLGTEAGHFAFFETGVLGPGGRAAWL
RSEPLPATPASCLRFWYHMGFPEHFYKGELKVLLHSAQQQLAVWGAGGHRHQWLEAQVE
VASAKEFQIVFEATLGGQPALGPIALDDVEYLAGQHCCQPAPSPGNTAAPGSVPAVVGSA
LLLLMLLVLLGLGRRWLQKKGSCPFQSNTEATAPGFDNILFNADGVTLPASVTS DP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**Transmembrane domain:**

Amino acids 1075-1092

**N-glycosylation sites:**

Amino acids 203-207;281-285;339-343;756-760

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 514-518;1100-1104

**N-myristoylation sites:**

Amino acids 32-38;55-61;61-67;106-112;116-122;336-342;350-356;409-415;  
523-529;540-546;678-684;707-713;791-797;870-876;921-927;  
937-943;954-960;1036-1042;1071-1077

**Amidation site:**

Amino acids 1093-1097

**Cell attachment sequence:**

Amino acids 191-194

523/615

**FIGURE 519**

GCAGGGGAGCTCCGAGTGTCCACAGGAAGGGAACTATCAGCTCCTGGCATCTGTAAGG**ATG**CT  
GTCCATGCTGAGGACAATGACCAGACTCTGCTTCCTGTTATTCTTCTCTGTGGCCACCAGTGG  
GTGCAGTGCAGCAGCAGCCTCTTCTCTTGAGATGCTCTCGAGGGAATTCGAAACCTGTGCCTT  
CTCCTTTTCTTCCCTGCCTAGAAAGCTGCAAAGAAATCAAGGAACGCTGCCATAGTGCAGGTGA  
TGGCCTGTATTTTCTCCGCACCAAGAATGGTGTGTCTACCAGACCTTCTGTGACATGACTTC  
TGGGGGTGGCGGCTGGACCCTGGTGGCCAGCGTGCACGAGAATGACATGCGTGGGAAGTGCAC  
GGTGGGTGATCGCTGGTCCAGTCAGCAGGGCAACAAAGCAGACTACCCAGAGGGGGATGGCAA  
CTGGGCCAACTACAACACCTTTGGATCTGCAGAGGCGGCCACGAGCGATGACTACAAGAACCC  
TGGCTACTACGACATCCAGGCCAAGGACCTGGGCATCTGGCATGTGCCCAACAAGTCCCCCAT  
GCAGCATTGGAGAAACAGCGCCCTGCTGAGGTACCGCACCAACACTGGCTTCCTCCAGAGACT  
GGGACATAATCTGTTTGGCATCTACCAGAAATACCCAGTGAAATACAGATCAGGGAAATGTTG  
GAATGACAATGGCCCAGCCATACCTGTGGTCTATGACTTTGGTGATGCTAAGAAGACTGCATC  
TTATTACTCACCGTATGGTCAACGGGAATTTGTTGCAGGATTCGTTTCAGTTCGGGGTGTTTAA  
TAACGAGAGAGCAGCCAACGCCCTTTGTGCTGGGATAAAAGTTACTGGCTGTAACTGAGCA  
TCACTGCATCGGTGGAGGAGGGTTCTTCCACAGGGCAAACCCCGTCAGTGTGGGGACTTCTC  
CGCCTTTGACTGGGATGGATATGGAATCACGTTAAGAGCAGCTGCAGTCGGGAGATAACGGA  
GGCGGCTGTACTCTTGTTCATAGAT**TGA**GACAGAGCTCTGCGGTGTCAGGGCGAGAACCCATC  
TTCCAACCCCGCTATTTGGAGACGAAAAACTGGAATTCTAACAAGGAGGAGAGGAGACTAA  
ATCACATCAATTTGCA

524/615

**FIGURE 520**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108701
><subunit 1 of 1, 325 aa, 1 stop
><MW: 36212, pI: 8.68, NX(S/T): 1
MLSMLRTMTRLCLLFFSVATSGCSAAAASSLEMLSREFETCAFSFSSLPRSCKEIKERC
HSAGDGLYFLRTKNGVVYQTFCDMTSGGGGWTLVASVHENDMRGKCTVGDRWSSQQGNKA
DYPEGDGNWANYNTFGSAEAATSDDYKNPGYYDIQAKDLGIWHVPNKSPMQHWRNSALLR
YRTNTGFLQRLGHNLFGIYQKYPVKYRSGKCWNDNGPAIPVVYDFGDAKKTASYYSYGQ
REFVAGFVQFRVFNNERAANALCAGIKVTGCNTEHHCIGGGGFFPQGKPRQCGDFSADFWD
DGYGTHVKSSCSREITEAAVLLFYR
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-26

**Glycosaminoglycan attachment site:**

Amino acids 86-90

**N-myristoylation sites:**

Amino acids 23-29;88-94;127-133;136-142;265-271

525/615

**FIGURE 521**

GATCAGTGTGTGAGGGAAGTCCATCATGAGGTCTGACAAGTCAGCTTTGGTATTTCTGCTCCTGCAGCTCTTCT  
GTGTTGGCTGTGGATTCTGTGGGAAAGTCCTGGTGTGGCCCTGTGACATGAGCCATTGGCTTAATGTCAAGGTCA  
TTCTAGAAGAGCTCATAGTGAGAGGCCATGAGGTAACAGTATTGACTCACTCAAAGCCTTCGTAAATTGACTACA  
GGAAGCCTTCTGCATTGAAATTTGAGGTGGTCCATATGCCACAGGACAGAACAGAAGAAAATGAAATATTTGTTG  
ACCTAGCTCTGAATGTCTTGCCAGGCTTATCAACCTGGCAATCAGTTATAAAATTAATGATTTTTTTGTTGAAA  
TAAGAGGAACTTTAAAAATGATGTGTGAGAGCTTTATCTACAATCAGACGCTTATGAAGAAGCTACAGGAAACCA  
ACTACGATGTAATGCTTATAGACCCTGTGATTCCCTGTGGAGACCTGATGGCTGAGTTGCTTGCAGTCCCTTTTG  
TGCTCACACTTAGAATTTCTGTAGGAGGCAATATGGAGCGAAGCTGTGGGAACTTCCAGCTCCACTTTCCTATG  
TACCTGTGCCTATGACAGGACTAACAGACAGAATGACCTTTCTGGAAAGAGTAAAAAATTCATGCTTTCAGTTT  
TGTTCCACTTCTGGATTACGATTACGACTATCATTTTTGGGAAGAGTTTATAGTAAGGCATTAGGAAGGCCCA  
CTACATTATGTGAGACTGTGGGAAAAGCTGAGATATGGCTAATACGAACATATTGGGATTTTGAATTTCCCTAAC  
CATACCAACCTAACTTTGAGTTTGTGGAGGATTGCACTGTAAACCTGCCAAAGCTTTGCCTAAGGAAATGGAAA  
ATTTTGTCCAGAGTTCCAGGGGAAGATGGTATTGTGGTGTCTCTGGGGTCACTGTTTCAAAATGTTACAGAAG  
AAAAGGCTAATATCATTGCTTCAGCCCTTGCCAGATCCCACAGAAGGTGTTATGGAGGTACAAAGGAAAAAAC  
CATCCACATTAGGAGCCAATACTCGGCTGTATGATTGGATACCCAGAAATGATCTTCTTGGTCATCCCAAAACCA  
AAGCTTTTATCACTCATGGTGGAATGAATGGGATCTATGAAGCTATTTACCATGGGGTCCCTATGGTGGGAGTTC  
CCATATTTGGTGATCAGCTTGATAACATAGCTCACATGAAGGCCAAAGGAGCAGCTGTAGAAATAAACTTCAAAA  
CTATGACAAGCGAAGATTTACTGAGGGCTTTGAGAACAGTCATTACCGATTCTCTTATAAAGAGAATGCTATGA  
GATTATCAAGAATTCACCATGATCAACCTGTAAAGCCCCCTAGATCGAGCAGTCTTCTGGATCGAGTTTGTCTATGC  
GCCACAAAGGAGCCAAGCACCTGCGATCAGCTGCCCATGACCTCACCTGGTTCCAGCACTACTCTATAGATGTGA  
TTGGGTTCTCTGCTGACCTGTGTGGCAACTGCTATATTCTTGTTCACAAAATGTTTTTTATTTTCTCTGCAAAAAT  
TTAATAAACTAGAAAGATAGAAAAGAGGGAATAGATCTTTCCAAATTCAGAAAAGACCTGATGGGGTAATCCTG  
TTAATTCAGCCACATAGAATTTGGTGAAAACCTTGCTATTTTCATATTATCTATTCTGTTATTTTATCTTAGCT  
ATATAGCCTAGAATTCATGATCATGAGGTTGTGAGTATATCTCATTCTTTTCGTTGATTTTTCCTAGGTGTCTTT  
ACTCTCTTCTCTCACTTTGTGACACAAGGACATGAATACATCTAAATTTTCTTATTTCTGATATGACTGTTTTGA  
TGATGTCACTTACTTCTATAACCTTAAGTGATAGGGTGACATGCAATATGATTATTCCTGGTGTGCGCCCAACAC  
ATGGATATAAAGAGGTAAAAAATTAATTCACAAAATTCAGTAAACCACACAAATCAGGTAAGTGTCTATGA  
GATTAGCTGGCTATGAGAAACATAATGATGTTTCTTTTTCAATTTAAATAAGCCTTTCTACATAGCCAGCATCAG  
TGATCTCAGAAAATAAATTGCTAATAATGATGACATGGCATTATGCTTAGAAAAGTTTGTGTATTTCCATAGAC  
CTCATCTAGATGTGATGGCCTACATTTCTGCCATCACTCAACCAATACTTTTTTCTGTTTTCTTGATGATAAAAA  
GACCTTTCTCATGATTGCCATCAAATAACAAAAGAACTATTTTTTTCTCACATAGAGAACATGTCAGTAAGAT  
ATTCAGGTGAACAGATATTTTTGGGATTAGTAACCTATTTGAAATATGTGGTGATAATTACTGAGTTTATAAAAT  
TTATTTGATAGTACACTTAAAGAAGATTTATATGTTTATTCTTTAAAAATGATGAATACTCATAATTCCTATCTC  
TATAATCAAAAGTATAATTTACTGTAGAAAAATAAGAGATGCTTGTTCTGAAAGTAAGATCAGTGAAGTGCCTT  
TCAGTCTCAATCTTTGAGAAATGTAAATTCATCAAATAATGCTTACATAGTAAAAATTTAAGGTATTAGAAAAC  
CTGCATAACAAATAGTATTATATATTAATATTTTGATATGTAAAGCTCTACACAAAGCTAAATATAGTGTAATA  
ATGTTTACACTAGTAAGCAAATATGTTAATCTTCTCATTTTTTTTACTGTGATATAATCTTAGTGATATGCCTATT  
AATAGTTTTAAATAAATAAATTGGCTTATCTGGCTTTTTGAAAATTTTGAAATCTTACAGATGTTGATTAGGTA  
TATCTACAAATTAATTTCAATTTTAAATGATGATATAAAAAATAAATAAGTATTTTTCTTGTGTATGTATACA  
ATAAATAAATAAATAATGTTTACTGTTTTGAAAGTTTCTTAAGTTTTA



526/615

**FIGURE 522**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108720
><subunit 1 of 1, 527 aa, 1 stop
><MW: 60284, pI: 8.31, NX(S/T): 3
MRSDKSALVFLLLQLFCVGCFCGKVLVWPCDMSHWLNVKVILEELIVRGHEVTVLTHSK
PSLIDYRKPSALKFEVVHMPQDRTEENEIFVDLALNVLPGLSTWQSVIKLNDFFVEIRGT
LKMMCESFIYNQTLMKKLQETNYDVMLIDPVI PCGDLMAELLAVPFVLTLLRISVGGNMER
SCGKLPAPLSYVPVPM TGLTDRMTFLERVKNSMLSVLFHFWIQDYDYHFWEEFYKALGR
PTTLCETVGKAEIWLIRTYWDFEFPQPYQPNFEFVGGLHCKPAKALPKEMENFVQSSGED
GIVVFSLSLGFQNVTEEKANIIASALAQIPQKVLWRYKGKKPSTLGANTRYDWIPQNDL
LGHPKTKAFITHGGMNGIYEAIYHGVPVMGVPIFGDQLDNIAHMKAKGA AVEINFKTMTS
EDLLRALRTVITDSSYKENAMRLSRIHHDQPVKPLDRAVFWIEFVMRHKGAKHLRSA AHD
LTWFQHYSIDVIGFLLTCVATAIFLFTKCF LFSCQKFNKTRKIEKRE
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-21

**Transmembrane domain:**

Amino acids 489-510

**N-glycosylation sites:**

Amino acids 131-135;313-317;518-522

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 67-71;340-344

**Tyrosine kinase phosphorylation sites:**

Amino acids 122-131;136-144

**N-myristoylation sites:**

Amino acids 19-25;276-282;373-379;377-383

**Amidation site:**

Amino acids 338-342

527/615

**FIGURE 523**

GGCTGCGGGGTGCGGCACGGAAG**ATG**CACGCGAGGCTCCTGGGGCTCTCGGCCCTGCTGCAGGC  
GGCCGAACAGAGCGCGCGCCTTTACACCGTGGCTTACTACTTCACCACAGGACGGCTTCTGTG  
GGGGTGGCTGGCCCTTGCTGTCCTCCTGCCCGGGTTCTTGGTCCAGGCCCTGAGCTACCTGTG  
GTTCCGAGCAGACGGGCATCCAGGGCATTGCTCCTTGGTGATGCTGCACCTCCTACAGCTTGG  
TGTTTGAAGCGGCACTGGGACGCTGCACTGACCAGTCTGCAGAAGGAACCTGGAGGCTCCCCA  
CCGAGGCTGGCTGCAGCTGCAGGAGGCCGACCTGTCGGGCCCTTCGACTCTTGGAGGCCCTGCT  
GCAGACTGGGCCCCACCTGCTGCTTCAGACATATGTTTTTCTAGCCTCAGACTTCACAGATATT  
GTGCCAGGGGTGAGCACCTGTCTTCTGGTCCCTCACTCTCCTGGGCACTGGTGTCTTACACT  
CGCTTCATGGGCTTCATGAAGCCAGGCCACCTGGCCATGCCATGGGCCCGCCCTCTTCTGCCAG  
CAGCTCTGGAGGATGGGCATGTTGGGAACCCGCGTGCTGAGTCTGGTTCTGTTCTACAAAGCC  
TACCACTTTTGGGTTTTTGTGGTTGCAGGTGCCCACTGGCTGGTGATGACATTCTGGCTTGTC  
GCCCAGCAGAGTGACATCATCGACAGCACCTGCCACTGGAGGCTGTTCAACCTGCTCGTGGGG  
GCCGTGTACATCCTCTGCTACCTCAGCTTCTGGGACAGCCCTTCTAGAAATAGGATGGTCACG  
TTCTACATGGTCATGCTGTTGGAGAACATCATCCTGTTGCTGTTGGCCACCGACTTCTCCAG  
GGGGCATCGTGGACCAGCCTGCAGACCATAGCTGGGGTCTGTCTGGATTCTGATTGGCAGT  
GTCTCACTGGTAATTTATTACAGCCTGCTGCATCCAAATCCACAGACATCTGGCAGGGCTGC  
CTAAGGAAGTCCTGTGGCATTGCAGGAGGTGATAAAACAGAGAGAAGAGATTCTCCCCGGGCC  
ACAGATCTAGCTGGGAAGAGAACCGAGAGCTCAGGCTCATGCCAAGGGGCAAGTTATGAACCA  
ACCATTTTAGGGAAGCCCCCTACCCCTGAGCAGGTCCCCCAGAGGCTGGGCTGGGGACCCAG  
GTTGCTGTGGAGGACTCTTTCCTCAGTCATCACCCTGGCTGTGGGTGAACTTGCCCTAAAA  
ACAGGAAATGTGTCTAAGATCAATGCCGCCTTTGGAGATAACAGTCCTGCCTATTGTCCACCT  
GCATGGGGGTTGAGTCAACAGGACTACCTGCAGAGAAAGGCCTTGTCTGCCAGCAAGAGCTC  
CCATCCTCATCCCGTGACCCCTCAACCTTAGAGAACAGCTCTGCGTTTGAAGGTGTCCCTAAA  
GCAGAGGCCGACCCATTGGAAACCTCAAGTTACGTATCTTTTGCCAGCGATCAGCAGGATGAA  
GCACCTACCCAGAACCCAGCAGCCACGCGAGGGGGAGGGCACCCCAAAGGAAGGAGCTGACGCT  
GTTTCTGGGACACAGGGGAAGGGGACAGGTGGGCAGCAGAGAGGAGGGGAAGGACAGCAGAGT  
TCCACGTTGTACTTCAGCGCCACTGCAGAAGTGGCCACATCCTCACAACAAGAAGGCAGCCCA  
GCTACTCTGCAAACGGCCCACTCTGGAAGGAGGCTGGGAAAGAGCAGCCCTGCCAGCCTGCA  
TCGCCCCACCCAGTGGGCTTGGCGCCCTTCCCCGACACCATGGCCGACATTAGCCCCATCCTA  
GGCACAGGCCCATGTAGAGGCTTCTGCCCCAGTGCAGGCTTCCCTGGAAGAACCCTCAGTATC  
TCAGAGCTAGAGGAGCCGCTGGAGCCCCAAAGGGAGCTAAGTCACCATGCAGCTGTTGGTGTG  
TGGGTGTCATTGCCACAGCTGAGGACTGCCCATGAGCCCTGCCTCACGTCCACCCCTAAGTCT  
GAGTCTATCCAAACGGACTGCAGCTGCAGGGAACAGATGAAGCAAGAGCCGAGTTTTTTCATC  
**TGA**CCACAGTCATGGTGGGATAAGACAACAGGCTGACAAACCAAGCTGGCCATTTGGTACCGT  
GAGAAAGGAAATCCCACTTCTGACACCTGTGTCTTGGGCACATCACTGTCACCTCTGAATCT  
CCATCTGCATCCCTGAAAAATGAAGAAACAGGGCTGGATGATTTTGCAGGTCCAATGCAAACA  
TCACAGACCCCAACCCATGCATAGGAGAGACTCTAACATACTTTAGAGGAGGAGAAAGAGATT  
CAGTCAAAATTGTCTGCTACCTTTTATGAGCTGTAGGTTCCCTTATTTTATCTTTTGTCTGTG  
GCTTCTAGGAAACACAAAGGTAAAACCCAGATTCTTATTTTATTTGAGGTTCTTGTTACAATT  
AGCTTTGCCTCACATTTAGCGGTTATGAATCTCATTTTAAATATATTCTAACTGTATTATGTTA  
TGAAATCTCTTGGTAAGATAATTTGCATGCTTTCTGGGAGTAGGTAAGGCCTGTGTGCTTGTA  
ATAACTAACATAACTGAAAGTGCAAATGTCA

528/615

**FIGURE 524**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108726
><subunit 1 of 1, 686 aa, 1 stop
><MW: 74981, pI: 6.60, NX(S/T): 2
MHARLLGLSALLQAAEQSARLYTVAYYFTTGRLLWGWLALAVLLPGFLVQALSYLWFRAD
GHPGHCSLVMHLHLLQLGVWKRHWDAALTSLQKELEAPHRGWLQLQEADLSALRLLEALLQ
TGPHLLLQTYVFLASDFTDIVPGVSTLFSWSSLSWALVSYTREMGMKPGHLAMPWAALF
CQQLWRMGMLGTRVLSLVLFYKAYHFWVFVAGAHWLVMTFWLVAQQSDIIDSTCHWRLF
NLLVGAVYILCYLSFWDSPSRNRMVTFYVMMLLENIILLLLATDFLQGASWTSLQTIAGV
LSGFLIGSVSLVIYYSLHHPKSTDIWQGCLRKSCGIAGGDKTERRDSPRATDLAGKRTE
SGSCQGASYEPTILGKPPTPEQVPPEAGLGTQVAVEDSFLSHHHWLWVKLALKTGNVSKI
NAAFGDNSPAYCPPAWGLSQQDYLQRKALSAQQELPSSSRDPSTLENSSAFEGVPKAEAD
PLETSSYVSFASDQQDEAPTQNPAAATQGEPTKEGADAVSGTQGKGTGGQQRGEGQQSS
TLYFSATAEVATSSQQEGSPATLQTAHSGRRLGKSSPAQPASPHVGLAPFPDTMADISP
ILGTGPCRGCPSAGFPGRTLISIELEEPLEPKRELSHHAAGVWVSLPQLRTAHEPCLT
STPKSESIQTDCSCREQMKQEPSFFI
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-17

**Transmembrane domains:**

Amino acids 35-50;269-287;293-313

**N-glycosylation sites:**

Amino acids 416-420;467-471

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 344-348

**N-myristoylation sites:**Amino acids 188-194;288-294;299-305;335-341;338-344;362-368;  
390-396;473-479;529-535;536-542;558-564;603-609;  
643-649**Amidation sites:**

Amino acids 354-358;568-572

**Leucine zipper pattern:**

Amino acids 112-134

529/615

**FIGURE 525A**

AGTGCCCTCTGTCATTTAATCTCCACGGAGTCCTGAAGGTGCTTCCCAGGTTTGTCCCCATCACACAGATGAGGC  
AATCGTTCTGTTAAGACTGTCCTTGGGTCAATTATGTCAGCAATGAAATCGGTGCTGCCATTACTAAACCCTTACT  
GTGTGCTGGCATTGTGTATGCATGCATGTGTGTGCGTGACACGTTGTGTGTTTGCCTGTACATGTGCATGTGTGTG  
TTGTGTGCCTGTGTGTGCACATGTAGAAAGAAAGTGATGTGTGGGAATGGAGAATTCCAACCCAGGAGAGACTG  
TGCTGGGGCTGCCACGAGAAGTGGTGACACTTCGAGAAACGGGTTCCAAATGCACGTTGCCTTCAAGTCTCTG  
TGTGACCTTGGGCAAGTCACCTCAGCTCCCTGAGCTTTGTTTTTAAAAATATTTTTTAAATGTATAAAACCATGG  
ACCATTACATATGAAGAGAAATGTGTGTGCAACATTCAAGTAAATAATCACAAGGTGGAGGAGTGCCTGCTCAGA  
CCCAGAGCTGCACAGGGAACTCTCCAGAAGGCTGCAGGGCTTCCCTCCCAACCCTCCAACGGCCCACTTGCTG  
AGCCCGTGCTTATCTGTTCAAGTGGAATGCCATGTGCCAGGTTCCCTCAGAAGTAAAGCCACTAGGTAAGTGTGAC  
ACTCTGTGCCAGCTTGGTGCAAAGCCCCCTTGTGTGTGCTATCTTATTGACCTCCCAATAGCATGGTAAGGTCA  
GTACTGTCACTTCCCCCACTTTGAAGATGAGGAGAGCACAATCTAGATGGAAATGGAGGTCACGCAGTGGAAA  
TAGGATCCAGACAGATTAATCCAATCTCAAGCCTGAATTTCTCCATTCCACGCTACGCTTGAAGCTCAATCTCTC  
TTCTGGTTGATTCTCCCCACTTCCCCACCCCAAGATATATCCCATCGCTGCTTGGTGGACAGTAGCCATGACTG  
GGTTTTGGTAAAGGTTGCTGAATAATCAGGCTGCTGGTTAGTTTTACATTTACCTTTCCAGTGAAATGGGGC  
CCCATGAAAAAGGCAGCTCAAGTTGTAAATTAATCAAAGGAAGGACAGAAAGGCTTCTGTTTGCACCTACCCTA  
AGGATTTGGGGTAGACACTGGGAATTTACTAATTATGAATTCAGTGCTTTCCTTGCTGAAAGAGAGGGCGTGGAA  
TCAACGCTGAGTGAAGGCATCAAGTTAAGCTGCTAATTACTTCTGATCATGCAGAATAAAAGCTACGTCCTT  
GAAATACACCAGGCAGCTAAACATAATCTTTGCGTTTCCGTAGTGTGGTTAAGGAATCCAGATGTTACTGCAAT  
AACCCTCCATAAACAAAAGGAACACCCAGCTGTGAGAACTGGCTTCTCAGCATTCGTCCAGCAGAGGCTCTTC  
CGGGGCCAGCCCTGGAAGAACCCATCAGGGTCTGATGGTTGCCCTGTTTCAGCACAGCCCTTATTGGCAGGCAG  
ACGGCTACGGGCACAGCCACAGGCTGAAGGTGAGTCCAGCACACAACCTTCTGACAGTGAACAGGAGTAAACATG  
GGACCCACCCGAAACCTTTGCTGTTGACTTCTTAGCAAATGGAGGCAGCTCTAGGCTCTGGAGAGTTCGGGTAT  
AGGAGACCATGACTTGAGCAGACTGATATAAGTGAATGCAACATATTTAGATGGCACAACCTAATTTAGATTT  
ATCAGTGCTAATATAGAAAAAGCTAGTATTTATTGGGGCTTATTAGATTTTGTAGTCTGAATCCTCACAACCTACG  
AGGGGGTTGTTTTACAGACTATGATCTTGCATGATTTCCCAAGATGCTCATTAAGTATATGGTGAAGTAGA  
ATTTGAATACAGAAGACCTGGTCTGCTACTTCTGTGTTTCTATTTGGTTCAAACAGCCCTTCTCTTTCAA  
ACAACCTTCAGTGCAATTCATGGTTTTGGAAAATAAACTTGATTTTGAGATTCAGACAATAAGTGCATTTTTAATG  
TTTATTCTTTTATCTTGA AAAACTGATATATTTATGAAATGATATGTCTCACTCAGTGTCAACACTTCAAACAA  
CACAGACAGTACAATGACAAATGGAGATCAGCTCTAATCTCGGCCCAATTTAATGCATTGCTGAATATTTCTT  
CTGAACATAGTCCATCCCACACTGTCCCATGACACAAGACGCTCCAAGGGGCTGAAGATAGAGGGACTTCTGCAG  
TCAAGAGAGCTGGGAAACTCTTGGACAGTCACAATGTGCATTTGGGTATTAAAGGCTCTGCAAAGTCTGCACCA  
AATAAACCTTGGATTGGCTTGATCCAATGCCATGTTTCCAAAACCTACTTGCCCGTGGGACACCTTAGTCCATA  
ACACAGGTTGGCATTCTTCTAGAGAGTGTGCTGTGAAAAACACTGGTCTCACAGCACCGTGCATTTCATCCAGCA  
GGTATTTACCAAGCAGGGACTTTGGGCCAGGTCCGTGCTAGGCTCTGCAGGTGGACCAGCCAGCCCTGACCTCCA  
TGGTGTCTCTTCTCATGGGAGAGGCTGCACAGCAGTCATTGAGAAAACGAAGAAACACACAGGTACTTTCAGATG  
CTGATAATGACTACCATGTGCTAAAAGAGCTCCAGGTGTTGTCTGTTTTGAGACAATCTCTCGACAATGAGATA  
GAATGAACCATGCAAACCTTTGGGGGCTACGATGGTTTTAGGAAAGAGCTAGAGTGA AAAATCCTTTGACATATAT  
ACATACAAATAAAGATACATGTGTATAATTTTATGTAATTGACATCACATTATATGCTGGCACCTGCCCTTTTTTG  
CTTAATGAAAATGTCAAAGTAATATTACCAATCAATCAATATATTTACCATCGTGTTTAATTTCTGCAGAGCA  
CTTATTCAAGCACATACTGCATTTTTCTTAGTCACTCACTGGTGGGAATAGAAAGTTTTCTCAGGCATGCATTTT  
TCCCACTTCTGAAATGATCTCCTTAGAATAAATCCCCAAATGTATTGCTGGGTTCAAAGGCATGAACATTTTA  
CATTTTTATACAATAATGTCAAACCTACCTTCCGGAATGATGCTCACTTTACTTTCCCTCCAAGACTGTGTGAAA  
ATGCCCATTTTTCTAAATGCTTACTATGACTGGTTTTCCAACCTACATTTAATTTCTTGTTCATCTGATAGGCAAA  
AAATGATATTTAATTTTTATTTGATTGATAATGACCTTGAACGTGCCCATTAGCCTTTTGCATGTATTCTTTTAT  
GAAACATCTGTTTCTATCCTTCGTCAATGTTTCTTCTATGTTAACTTTTCTTATTGATTTGTTAGAGCACTT  
TGTATATTGTGAATATTAGCTTTTGTGATCAGAATTATGGAATGTTGTTGTCTTTAGTTTTGCTTATCAAGTTT  
AAAGCCATTACAGAGATGTTGTAATGTGTATGTTGTTAAATTTATGTCTTATTTTGTGTTTTCTGATGCTCATAT  
GTTTAGAATGGTCAAGCAATCCCACTTATGATTACATAAATTTCTCCATATTTGCTCTTAGCATTTTTTTTC  
TTTTCATGTAAACCTTTGTTCTTTTCAAGAAATTTACTTTCGGTAAGAAATGGGCTGGCTTCCAGTTTTATGTTT  
CCAAATGGTGATTGAGCTGTCCATTTCCATATTTCTCCCTTATTAGAAATGACCACTATATTATGTTCTAAAATA  
TCTGCGTACTTGTGTCCCTTCTATAATCTCAGTTCATCTTTTGTGCTATCTTTTGTGCTTTTCAAACCAC  
ACTGCTTTACTGAAGTGTATCTTATACATTTTTTAATCTCAGCAAGACAAGTTTCTCAATGCCACTCTTTT  
TCAGAGTTTTTCTGGTGGTTGTAAGATGTTTATTCTTCTGGATAAACTTTAGAATCACTCTTTTTGTCCAAGGTA  
AAATATATCCCACATTGAGATCATCTGAATATACAGACTAATTCAGGAAAAATGTATGCTTTATTGCAATGA  
GTCTTCTTATCCAATAAAAAAGATATGAATTTCCATGTATTGAAATCTTCACTGAGACTTATTTTTGGCTTTTCA

530/615

**FIGURE 525B**

CATGTCCTGCAAATGATTGTTAAATTTATTTTATAGGTATTTAGGGGAAATGATTTTCTAAAGTTTGTATTTTC  
TAGCTTGTTATAATTTACATATGAGATAGTCATTGTTGTATATTATTTATAACTGATCATATTACTGTATTTGTA  
TTGTTTTAATAGTTTTTCTATTATTTTGGGTTTTCTGGAATACAACCTTATTATCTACAAATTATGATTGTTTT  
GCCTTTTCCAATGTTCACTGTTTTATATTCTTGTCTGATTGCTTTGTTTCAGCACTTCTAGAATAAAGTCAT  
GCAATACTAATGA

531/615

**FIGURE 526**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108728
<subunit 1 of 1, 100 aa, 1 stop
<MW: 10922, pI: 8.81, NX(S/T): 0
MSAMKSVLPLLNPYCVLAFVYACMCVRAHVCVVYMCMCVLCACVCTCRKKVMCGNGEFQ
PRRRLCLGLPREVVTLRETGSKCTLPSSSLCDLGQVTSAP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-28

**N-myristoylation sites:**

Amino acids 80-86;94-100

**Prokaryotic membrane lipoprotein lipid attachment sites:**

Amino acids 13-24;34-45

532/615

**FIGURE 527**

GTGAGACTTCTTTCTTCATTGTGGCTAGCTTTGAAAAGACCCTCTGAACTTCCTAAAGATATC  
AAG**ATG**ATATCACCAGACTTGCCCTTTTTGACAATTGTCTTGATCATAGTTAGTTGGACAACCT  
TGTGGAGCACTAGCCATACTTCTTTCTTATCTTTACTATGTGTTTAAGGTTGTTTCATCTGCAA  
GCCAGCTTAACAACCTTTTAAGAATAGCCAGCCTGTGAATCCCAAACACTCTAGAAGAAGTGAA  
AAGAAATCCAATCATCATAAAGACTCCTCAATACACCATCTTCGTTTATCTGCCAACGATGCT  
GAAGATAGCCTTCGCATGCACAGTACTGTGATTAACTTACTAACATGGATTGTATTACTCAGC  
ATGCCTTCTCTAATTTATTGGCTAAAGAATCTTAGGTATTATTTTAACTTAATCCTGATCCA  
TGTAACCTTTGGCATTATCCTTATTCCGACTATGGCAATTCTTGGAATACTTACACTGTT  
TCAATAAAATCAAGTAAATTGTTGAAGACTACTTCACAATTTCCACTTCCTCTGGCTGTTGGT  
GTGATTGCTTTTGGGTCAGCACATTTATATAGGCTTCCATGCTTTGTCTTCATTCCTCTTTTA  
CTCCATGCATTATGCAACTTTATG**TAA**GATTGGACTTAAGGAATGATGAAGATAATTTATGTG  
TTTAGGGCCAGTGATAAGAGGGAACACACAGATCCATCAGTATGGACAGCAAGATCCTTTGGA  
GAAGACAAGTCTATTTTTACAATATTGAAAATAGGAAATTAGTTTTGTAATGTTTGAGGGAAG  
TAGTTGAAGCATGGTTTTGTTTTGTGGTGTGGAATCCATGTACTAATCATTTTTGAAAATTC  
ATGAAGGGATATATGGTGATCACTATCATTGAGGACTCCTGTGCATATAAAATAGTCTGTTTT  
ATCAACTGTAAA

533/615

**FIGURE 528**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108738
><subunit 1 of 1, 196 aa, 1 stop
><MW: 22225, pI: 9.90, NX(S/T): 0
MISPDLPFLTIVLIIVSWTTCGALAILLSYLYYVFKVVHLQASLTTFKNSQPVNPKHSRR
SEKKSNNHHKDSSIHHLRLSANDAEDSLRMHSTVINLLTWIVLLSMPSLIYWLKNLRYYYFK
LNPDPCKPLAFILIPMAILGNTYTVSIKSSKLLKTTSQFPLPLAVGVIAFGSAHLYRLP
CFVFIPLLLHALCNFM
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-25

**Transmembrane domains:**

Amino acids 91-108;128-143;167-186

**N-myristoylation site:**

Amino acids 141-147



534/615

**FIGURE 529**

GCGAGCCGGGTCCCACCATGGCCGCGAATTATTCCAGTACCAGTACCCGGAGAGAACATGTCA  
AAGTTAAAACCAGCTCCCAGCCAGGCTTCCTGGAACGGCTGAGCGAGACCTCGGGTGGGATGT  
TTGTGGGGCTCATGGCCTTCCTGCTCTCCTTCTACCTAATTTTCACCAATGAGGGCCGCGCAT  
TGAAGACGGCAACCTCATTGGCTGAGGGGCTCTCGCTTGTGGTGTCTCCCGACAGCATCCACA  
GTGTGGCTCCGGAGAATGAAGGAAGGCTGGTGCACATCATTGGCGCCTTACGGACATCCAAGCTT  
TTGTCTGATCCAAACTATGGGGTCCATCTTCCGGCTGTGAAACTGCGGAGGCACGTGGAGATG  
TACCAATGGGTAGAAACTGAGGAGTCCAGGGAGTACACCGAGGATGGGCAGGTGAAGAAGGAG  
ACGAGGTATTCTTACAACACTGAATGGAGGTGAGAAATCATCAACAGCAAAAACTTCGACCGA  
GAGATTGGCCACAAAAACCCAGTGCCATGGCAGTGGAGTCATTCATGGCAACAGCCCCCTTT  
GTCCAAATTGGCAGGTTTTTCTCTCTCGTCAGGCCTCATCGACAAAGTCGACAACTTCAAGTCC  
CTGAGCCTATCCAAGCTGGAGGACCCTCATGTGGACATCATTCGCCGTGGAGACTTTTTCTAC  
CACAGCGAAAATCCCAAGTATCCAGAGGTGGGAGACTTGCCTGTCTCCTTTTCTATGCTGGA  
CTGAGCGGCGATGACCCTGACCTGGGCCAGCTCACGTGGTCACTGTGATTGCCCCGCGAGCGG  
GGTGACCAGCTAGTCCCATTTCTCCACCAAGTCTGGGGATACCTTACTGCTCCTGCACCACGGG  
GACTTCTCAGCAGAGGAGGTGTTTCATAGAGAACTAAGGAGCAACTCCATGAAGACCTGGGGC  
CTGCGGGCAGCTGGCTGGATGGCCATGTTTATGGGCCTCAACCTTATGACACGGATCCTCTAC  
ACCTTGGTGGACTGGTTTTCTGTCTTCCGAGACCTGGTCAACATTGGCCTGAAAGCCTTTGCC  
TTCTGTGTGGCCACCTCGCTGACCCTGCTGACCGTGGCGGCTGGCTGGCTCTTCTACCGACCC  
CTGTGGGCCCTCCTCATTGCCGGCCTGGCCCTTGTGCCCATCCTTGTTGCTCGGACACGGGTG  
CCAGCCAAAAAGTTGGAGTGA~~AA~~AGACCCTGGCACCCGCCGACACCTGCGTGAGCCCTGAGG  
CTGGTTGTACAATGCCACGCCTGCCTGGCTGCTTTACCTGGGAGTGCTTTCGATGTGGGCA  
CCTGGGCTTCCTAGGGCTGCTTCTGAGTGGTTCTTTACGTGTTGTGTCCATAGCTTTAGTCT  
TCCTAAATAAGATCCACCCACACCTAAGTCACAGAATTTCTAAGTTCCCCAACTACTCTCACA  
CCTTTTTAAAGATAAAGTATGTTGTAACCAGGACGTCTTAAATGATTCTTTGTGTACCTTTTC  
TGTATATTTCAGAAACCGTTCTGTGCCTGCTGGGAGTAATTCTTTAGCAATTAAGTATTTGG  
TAGCTGAATAAGGGGTGAGAACTTCTGAAACCAGAGATCTGTAATCATCTCTATTGGCCTGGG  
GTGCCTGTGCTATAAATGAGTTTCTTACATGAAAAACACAGCCAGCCCAAGATGACTTATCT  
GGGTTTAGGATTCAATAGTATTCCTAACTGCTTATTACATGAGCAATTCATCAAATCTCCA  
AACTCTTAAAGGATGCTTTTCGGAAAACACGCTGTATACCTAGATGATGACTAAATGCAAAATC  
CTTGGGCTTTGGTTTTTTTTCTAGTAAGGATTTTAAATAACTGCCGACTTCAAAGTGTCTTA  
AAACGAAAGATAATGTTAAGAAAAATTTGAAAGCTTTGGAAAACCAATTTGTAATATCATTG  
TATTTTTTATTAAAAGTTTTGTAATAAATTTCTAAATTATCA

535/615

**FIGURE 530**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108743
><subunit 1 of 1, 400 aa, 1 stop
><MW: 44876, pI: 8.32, NX(S/T): 2
MAANYSSSTSTRREHVKVKTSSQPGFLERLSETSGGMFVGLMAFLLSFYLIIFTNEGRALKT
ATSLAEGLSLVSPDSIHSVAPENEGRLVHIIGALRTSKLLSDPNYGVHLPVKLRRHVE
MYQWVETEEESREYTEDGQVKKETRYSYNTEWRSEIINSKNFDREIGHKNPSAMAVESFMA
TAPFVQIGRFFLSSGLIDKVDNFKSLSLSKLEDPHVVDIIRRGDFFYHSENPKYPEVGDLR
VSFSYAGLSGDDPDLGPAHVVTVIARQRGDQLVPFSTKSGDTLLLLHHGDFSAAEEVFHRE
LRNSMKTWGLRAAGWMAMFMGLNLMTRILYTLVDWFPVFRDLVNIGLKAFACVATSLT
LLTVAAGWLFYRPLWALLIAGLALVPILVARTRVPAKKE
```

**Important features of the protein:****Transmembrane domains:**

Amino acids 34-53;365-388

**N-glycosylation site:**

Amino acids 4-8

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 140-144

**Tyrosine kinase phosphorylation sites:**

Amino acids 99-107;220-227

**N-myristoylation sites:**

Amino acids 35-41;93-99;310-316

**Cell attachment sequences:**

Amino acids 221-224;268-271

536/615

**FIGURE 531**

AAAAAAAAAAAAAAAAAAGAAGCTCTTATGCCAGGAACCTGGAATGGAGACCAAATATATATTG  
GTTATATCATAGTATCACAGGGTTACTTTGGCATTGTTGGGAACTTGAGAGAAATGGGCAATAA  
CTGTTACTTTAAAAGCTTGGGTGCTGTGATTCTGCCTTCAGCCTCAGCCACTTTTGTGGTGCT  
TTGCGTGGCATCAGTACCTCCACTGATTCTTCTGTCTTTCCTCTCTCTTCCCCCCTCTTT  
CCCTTCTGTTTTTCTCAGATCTAAGGGTTATAATGGAGGGGCAAACCTGCCTGGCTATTTCAGA  
TAAGACTTCACTGAGTGACTGTTTCAGCCCATGATTTACCCTGCAGTTTAACAGGCTCAGGAAT  
TAGGTCGCATCAGTTGAGCGCGGGTCACTTAGGCCTATAATCATCATCAGACGGCAATTAAAG  
GACCATTTCTGCCTTTTTCACTATTACATCCCCCGCCTGTAGCCAGCCTGCCATACAGTAGA  
TACTCAATAAATATTTGCTGAATGATAACCAATAA

537/615

**FIGURE 532**

</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108758  
<subunit 1 of 1, 100 aa, 1 stop  
<MW: 10316, pI: 8.52, NX(S/T): 0  
MGNNCYFKSLGAVILPSASATFVLCVASVPPLILLSFLSLFPPSFPSVFLRSKGYNGGA  
NCLAISDKTSLSDCSAHDLPCLTGS GIRSHQLSAGHLGL

**Important features of the protein:****Signal peptide:**

Amino acids 1-47

**N-myristoylation site:**

Amino acids 58-64

538/615

**FIGURE 533**

CGGGGTGTACGAAAGAGAAACCCGGAGGGCGCCGGGGGACTGGGCCGGGGTCTGCAGGGCTCAG  
CTGAGCCCATGAGCTCCCAGAGCTAACCCCTGAACACCCAGGCGGGCAAAGGGCTGATGTCGG  
TAGTCCCCATCCTGGAGGGGCAGGCTCTGCGCATCTGCTCCTGGC**ATG**GCGCTGCGGCACCTC  
GCCCTCCTGGCTGGCCTTCTCGTGGGAGTCGCCAGCAAGTCCATGGAGAACACGGCCCAGCTG  
CCCGAGTGCTGTGTGGATGTGGTGGGCGTCAACGCCAGCTGCCAGGCGCAAGTCTGTGTGGT  
CCAGGCTGTTACAGGCGCTGGAACGCGGACGGGAGCGCCAGCTGCGTCCGCTGTGGGAACGGA  
ACCTCCCAGCCTACAACGGCTCCGAGTGTAGAAGCTTTGCTGGCCCGGGTGCGCCATTCCCC  
ATGAACAGAAGCTCAGGGACCCCCGGGCGGCCACATCCTGGGGCTCCGCGCGTGGCCGCCTCC  
CTCTTCCTGGGCACGTTCTTCATTAGCTCCGGCCTCATCCTCTCCGTAGCTGGGTTCCTTCTAC  
CTCAAGCGCTCCAGTAAACTCCCCAGGGCCTGCTACAGAAGAAACAAAGCTCCGGCCCTGCAG  
CCTGGCGAAGCCGCTGCAATGATCCCCCGCCACAGTCTCAGACGTGGGGTCTGCAGGAAAG  
GAGGACCCACCACGACAGGGCAGACCCCCAATAACCTGCTCCTCCT**TGA**AGTCCAGCTCCACCC  
GAGGACAGACGCAGCCGGCCTCCGCCAGGGCCTCCTGAGCAGCCATCGCTTCAGTGGTGTGG  
GTCAGGCGGACCCAAGAGTCAGCCCGTACGGAAGCCGCGCTACGTGAGGCGGGAGCGGCCCT  
GGACAGGGCCACGGATCCCGCTGCCTTCCCGGGGGAGGCGCGTATCAGCAATGTCTGACCTGG  
AGGCCGAGACCACGCCACGCACTTGGCGGCAGGGACCCGGAGGCCGACCCCTTGGCGGGAACC  
AGCACAAAGTGTTGGCATCGCCCGGCGCCCGGGACAGTCTTGGGCACAGCCTCGGCTCTGGGT  
CCCTCCGCCTCCCAGCGACGGACGCCAAAGGGTCCCGGGCCGCCTGAGGCTCCTCCCCACCAC  
AGCCATCTCGTTTATCGGACCAGGAGCAGGCATCCATGAGACCTCAGAGCTTCAGATCGAGGC  
CTTGGGGGGTCCGGGGCCCCCCCCAGGAAACACGGTGAGGCCCCAGCGCTGCAGCCAAAGCTGG  
CACGATCTATGGGGCAGGTGCCGCTCTGCCTAGAAAAGCCAGGGGCTCTGCTGCCGTGCCCTC  
CAGAGCCCACAGCGGGCAGGACTCCTCCAGCACACCACACCCAGTGGCCCCGAGACCCCTCTG  
AGAACAGTGAGGCTGGTCTCGTGCCGTTCCAGCCGGTGCCCGGCCAGTGGGGAGGACACAGC  
CTAGGAACCAGCTGCCTGAGACCAGGGTGCCTCTGGGCTGTCTCCCGCGTGGCGGAGACCCC  
AAGCACGCAGCCACCCATTTCCGGAGCTGCAGGATAGAGCTTCCTCTTGATCTCTGTTTTTAAG  
CAGAAATTCATTGTGCAGAAAAGTCTCCAGAGCTCTGTGGCCCCGCTCGGATCCGCTGGACC  
CCCATGCCTGGCTGATCCCTGCCCACGTGGGGCAGGCCACATCTAACCCCCACAAGTCACTG  
CCTCACTGCACCTGCCAAGGCTGCCCTGGCGCTGAGTCTGGGGTCCCTCCCGGAGTTCCTGG  
GAGAAAGGCGCCGTCGTGGCCGCCTCCCGCACGCCAGGCCCGGGCTCCACCGTGGGTCTCAGA  
CGCCCTGCGGCACCGGCACCGTCTGCTTTAGCATGGGACCCCCATCTGAGGGGTGGCCTGGCC  
TTCGGGGTCCCCACGCTCCTTTGCGAAGTCCACTGTGGGTGCCATCATGGTCTCCGGGACCTG  
GGCCAGCGGGAACGTGGGGGCACTGGGTGTGCTGATATAAAGTCGGCATTACTCAAAAAAAAAA  
AAAAAAAAAAAAAAAAAAAAAAAAAAAAA

539/615

**FIGURE 534**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108765
><subunit 1 of 1, 189 aa, 1 stop
><MW: 19464, pI: 9.60, NX(S/T): 4
MALRHLALLAGLLVGVASKSMентаQLPECCVDVVGVNASCPGASLCGPGCYRRWNADGS
ASCVRCGNGTLPAYNGSECRSFAGPGAPFPMNRSSGTPGRPHPGAPRVAASLFLGTFFIS
SGLILSVAGFFYLKRSSKLPRACYRRNKAPALQPGEEAAMIPPPQSSDVGSAGKEDPPRQ
GRPPIPAAPP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-18

**Transmembrane domain:**

Amino acids 111-129

**N-glycosylation sites:**

Amino acids 38-42;68-72;75-79;92-96

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 134-138

**N-myristoylation sites:**

Amino acids 11-17;36-42;43-49;59-65;69-75;122-128

540/615

**FIGURE 535**

TGGATCTGCGGGAATGTGGGCTGGAGAGGTCCTGCCGTGGTACCAGCCTCCAGCCTGCCCCCA  
GGAAGTCCCCCTGACCCAGGCGCGCCCGCTGCTCGGTGGCAGGAGGGCCGGCGGAGCGCC**ATGG**  
CCTGCATCCTGAAGAGAAAGTCTGTGATTGCTGTGAGCTTCATAGCAGCGTTCCTTTTCCTGC  
TGGTTGTGCGTCTTGTAAATGAAGTGAATTTCCCATTTGCTACTAAACTGCTTTGGACAACCTG  
GTACAAAGTGGATAACCATTTCTCCTACACATACAGGCGGCCCTTCGAACTCACTATGGATACA  
TAAATGTGAAGACACAAGAGCCTTTGCAACTGGACTGTGACCTTTGTGCCATAGTGTCAAACCT  
CAGGTCAGATGGTTGGCCAGAAGGTGGGAAATGAGATAGATCGATCCTCCTGCATTTGGAGAA  
TGAACAATGCCCCACCAAAGGTTATGAAGAAGATGTCGGCCGCATGACCATGATTCGAGTTG  
TGTCCCATACACAGCGTTCCTCTTTTGCTAAAAAACCCCTGATTATTTTTTTCAAGGAAGCGAATA  
CTACTATTTATGTTATTTGGGGACCTTTCCGCAATATGAGGAAAGATGGCAATGGCATCGTTT  
ACAACATGTTGAAAAAGACAGTTGGTATCTATCCGAATGCCCAAATATACGTGACCACAGAGA  
AGCGCATGAGTTACTGTGATGGAGTTTTTAAGAAGGAAACTGGGAAGGACAGTACAGAG**TGAC**  
CATGCAGTGTTGATTGATCGAACAGCAACCACCACATACATGTCCTGCCCCACCACAAAAGGA  
AGGAAGGAATAAAAGAAAGAAAGAAAGAAACAAACAAACAAACAAACAAACTAAGCAAGACA  
AAACAAATACCCATGTCAGTGGTTCAAAGATTAAGATTGTGGCTTTGTGTAAAGTTCTTTCCC  
TTTGTAGACTTGCTGCATAATTATTCAGGTATGATGGTTACAGTTTTTAAAAAGGAAGGGAAA  
TTGTGGTATGTGGTATGTAAATATTTTAAATGTTGTCTCTCTGTTTTGATCAGTTTTTTGTTT  
TATTCAATTTGTCTTTATTAAATCTTATCAAAGCA

541/615

**FIGURE 536**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108783
><subunit 1 of 1, 210 aa, 1 stop
><MW: 24022, pI: 9.51, NX(S/T): 1
MACILKRKSVIAVSFIAAFLFLLVVRLVNEVNFPLLLNCFGQPGTKWIPFSYTYRRPLRT
HYGYINVKTQEPLQLDCDLCAIVSNSGQMVGQKVGNEIDRSSCIWRMNNAPTKGYEEDVG
RMTMIRVVSHTSVPDLLKNPDYFFKEANTTIYVIWGPFRNMRKDGNGIVYNMLKKTVGIY
PNAQIYVTTEKRMSYCDGVFKKETGKDSTE
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-27

**N-glycosylation sites:**

Amino acids 148-152

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 6-10;191-195;201-205

**N-myristoylation sites:**

Amino acids 41-47;87-93;91-97;167-173;178-184



542/615

**FIGURE 537**

GTTTTATTGACAATACATGCATCATATCTTTTGACTTTGAAGGATATCTCATGTCAAAGGAAT  
CAAGTTATGATTTATAGAGGATTCAGCTGGAATACCTTGTGGGTGCTGGCTGAGGGTGGCAAA  
ACGCCTACCGAGAC**ATGA**AGGTTTTAGCCACTAGTTTTGTCTTGGGAGCCTGGGGTTGGCCT  
TCTACCTGCCTTTGGTGGTGACTACACCTAAAACACTGGCCATCCCTGAGAAGCTGCAAGAAG  
CTGTGGGGAAAGTTATCATCAATGCCACAACCTGTACTGTCACCTGTGGCCTTGGCTATAAGG  
AGGAGACCGTCTGTGAGGTGGGCCCTGATGGAGTGAGAAGGAAATGTCAGACTCAGCGCTTAGAA  
TGTCTGACCAACTGGATCTGTGGGATGCTCCATTTACCATTCTCATTGGCAAGGAATTTGAG  
CTTAGCTGTCTGAGTTCAGACATCTTGGAGTTTGGACAGGAAGCTTTCGGGTTACCTGGAGA  
CTTGCTCGAGGTGTCATCTCCACTGACGATGAGGTCTTCAAACCCTTTCAAGCCAACCTCCAC  
TTTGTGAAGTTTAAATATGCTCAGGAGTATGACTCTGGGACATATCGCTGTGATGTGCAGCTG  
GTAAAAAACTTGAGACTTGTCAAGAGGCTCTATTTTGGGTTGAGGGTCCTTCCTCCTAACTTG  
GTGAATCTGAATTTCCATCAGTCACTTACTGAGGATCAGAAGTTAATAGATGAGGGATTGGAA  
GTTAATCTGGACAGCTACTCCAAGCCTCACCACCCAAAGTGGAAAAGAAGGTGGCGTCAGCC  
TTGGGAATAGGAATTGCCATTGGAGTGGTTGGTGGCGTGTTGGTGAGGATTGTCTCTGTGCG  
CTAAGGGGGGGCCTGCAGCAG**TGA**CAGCTTCAAGAACTTAACAGCCTTGCTCCTGAAGAACTG  
GCTGCCCAGGAAGCCAAGCTAGCTTTTTAGGGGAGTGTTCCAGCTGCTGGTAGTGGATCAGCT  
TAGAGGGAACTCTCCACAGCCAAAAGAATGAGTGGGAGAAATGGAGGGGACAATCTCCTGGG  
AGCTATGCGCAGTAACCTAACTTCCTTATGTCCCATGGATCTCTCCTGATCTTCCCTGCCA  
TTGGGTACCCAGGAACTGCAAGCATTGCCTGTGTTCTTGGGAAGAGTTCTAAGAAGCTTGCA  
TTCATTTTCTACCCTTTATGACTTGGATGCCTCCCCACCTCCATTTCCCCTCTTCTGAGCTGT  
GTATTCATGTAGAGGGATGTATTCAGCCTTTTTAGTGAACATTTTTTTTCAATAAAAGTAATT  
CACAGTAA

543/615

**FIGURE 538**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108789  
><subunit 1 of 1, 255 aa, 1 stop  
><MW: 28440, pI: 8.92, NX(S/T): 1  
MKVLATSFVLGSLGLAFYLPVVVTPKTLAIPEKLQEAVGKVIINATTCTVTCGLGYKEE  
TVCEVGPDGVRRKCQTQRLECLTNWICGMLHFTILIGKEFELSCLSSDILEFGQEAFT  
WRLARGVISTDDEVFKPFQANSHFVKFYAQEYDSGTYRCDVQLVKNLRLVKRLYFGLRV  
LPPNLVNLNFHQSLTEDQKLIDEGLEVNLDSYKPHHPKWKKKVASALGIGIAIGVVGGV  
LVRIVLCALRGGLQQ

**Important features of the protein:****Signal peptide:**

Amino acids 1-30

**Transmembrane domain:**

Amino acids 225-244

**N-glycosylation site:**

Amino acids 45-49

**N-myristoylation sites:**

Amino acids 126-132;156-162;204-210;229-235;231-237;235-241

544/615

**FIGURE 539**

GCGCTCATCACTGGCTGGGGACAGAGCCGGGCACCAAGGAGCGACAGGATCCCGAAGAGAGAG  
AGAGAAGGCAGCGAGGGAAGGAGGACCCCGGCAGGCAGCAGCATGAAATTCAGCCCAGCGCAC  
TACCTGCTGCCTCTCCTGCCTGCGCTGGTCCTCAGCACCCAGACAGGACTATGAAGAGCTAGAA  
AAGCAGCTGAAAGAAGTCTTTAAGGAGCGAAGCACCATTTCTTCGTCAGCTGACAAAGACATCA  
AGAGAACTTGATGGAATTAAAGTCAATCTTCAGTCCTTAAAAAACGATGAGCAGTCTGCCAAA  
ACTGATGTTTCAGAACTTCTGGAATTAGGACAGAAACAAAGAGAAGAAATGAAGTCTCTTCAG  
GAGGCCCTGCAAAATCAGCTTAAGGAGACATCAGAGAAAGCAGAAAAACACCAGGCTACTATT  
AATTTTTTAAAGACTGAAGTTGAAAGAAAGAGCAAAATGATCCGAGACCTCCAGAATGAGGAT  
TCAAGGAAGAGACCAAGAGATCTCCAGTGGAAGATAGTCTCCATGAGGACCATGTCAATATAC  
TTATTGATGTATCTCTAGTACCTAGAATAGTGGAGATTTATATTAGATACAAAATAAATATGT  
GTGGAATTAATTAATAA

545/615

**FIGURE 540**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108806
><subunit 1 of 1, 159 aa, 1 stop
><MW: 18865, pI: 9.76, NX(S/T): 0
MKFSPAHYLLPLLPAVLSTRQDYEELEKQLKEVFKERSTILRQLTKTSRELDGIKVNLO
SLKNDEQSAKTDVQKLELGQKQREEMKSLQEALQNQLKETSEKAEKHQATINFLKTEVE
RKSKMIRDLQNEDSRKRPRDLQWKIVSMRTMSIYLLMYL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-22

**N-myristoylation site:**

Amino acids 54-60



547/615

**FIGURE 542**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA108936
><subunit 1 of 1, 178 aa, 1 stop
><MW: 19472, pI: 5.71, NX(S/T): 0
MSPSGRLCLLTIVGLILPTRGQTLKDTTSSSSADSTIMDIQVPTRAPDAVYTELQPTSPT
PTWPADETPQPQTQTQQLEGTDGPLVTDPEETHKSTKAAHPTDDTTTLSERPSPSTDVQTD
PQTLKPSGFHEDDPFFYDEHTLRKRGLLVA AVLFITGIIILTS GKCRQLSRLCRNRCR
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-21

**Transmembrane domain:**

Amino acids 147-162

**Tyrosine kinase phosphorylation site:**

Amino acids 45-52

**N-myristoylation site:**

Amino acids 146-152

548/615

**FIGURE 543**

CGGCTCGAGGTGAGAAGGAACTGCAAGAGTGGGGCAGAGAACCAGAGTGTGAGAGCAAACC  
TCCTCTATCTGCACATCCTGGGGACGAACCGGGCAGCCGGAGAGCTGCGGCCGGCCAGTCCC  
GCTCCGCCCTTTGAAGGGTAAAACCCAAGGCGGGGCCCTTGGTTCTGGCAGAAGGGACGCT**ATGA**  
CCGCAGAATTCCTCTCCCTGCTTTGCCTCGGGCTGTGTCTGGGCTACGAAGATGAGAAAAAGA  
ATGAGAAACCGCCCAAGCCCTCCCTCCACGCCTGGCCCAGCTCGGTGGTTGAAGCCGAGAGCA  
ATGTGACCCTGAAGTGTGAGGCTCATTCCAGAATGTGACATTTGTGCTGCGCAAGGTGAACG  
ACTCTGGGTACAAGCAGGAACAGAGCTCGGCAGAAAACGAAGCTGAATCCCCCTTCACGGACC  
TGAAGCCTAAGGATGCTGGGAGGTACTTTTGTGCCTACAAGACAACAGCCTCCCATGAGTGGT  
CAGAAAGCAGTGAACACTTGCAGCTGGTGGTCACAGATAAACACGATGAACTTGAAGCTCCCT  
CAATGAAAACAGACACCAGAACCATCTTTGTGCGCCATCTTCAGCTGCATCTCCATCCTTCTCC  
TCTTCCTCTCAGTCTTCATCATCTACAGATGCAGCCAGCACGGTTCATCATCTGAGGAATCCA  
CCAAGAGAACCAGCCATTCCAACTTCCGGAGCARGAGGCTGCCGAGGCAGATTTATCCAATA  
TGGAAAGGGTATCTCTCTCGACGGCAGACCCCAAGGAGTGACCTATGCTGAGCTAAGCACCA  
GCGCCCTGTCTGAGGCAGCTTCAGACACCACCCAGGAGCCCCCAGGATCTCATGAATATGCGG  
CACTGAAAGTGT**TAG**CAAGAAGACAGCCCTGGCCACTAAAGGAGGGGGGATCGTGCTGGCCAAG  
GTTATCGGAAATCTGGAGATGCAGATACTGTGTTTCCTTGCTCTTCGTCCATATCAATAAAAT  
TAAGTTTCTCGTCTTA

549/615

**FIGURE 544**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119510
><subunit 1 of 1, 236 aa, 1 stop
><MW: 26079, pI: 5.05, NX(S/T): 3
MTAEFLSLLCLGLCLGYEDEKKNEKPPKPSLHAWPSSVVEAESNVTLKCOAHSQNVTFVL
RKVNDSGYKQEQSSAENEAEFPFTDLKPKDAGRYFCAYKTTASHEWSESSEHLQLVVTDK
HDELEAPSMKTDTRTIFVAIFSCISILLFLSVFIIYRCSQHGSSEESTKRTSHSKLPE
QEAAEADLSNMERVSLSTADPQGVTYAELSTSALSEAASDTTQEPPGSHEYAALKV
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-16

**Transmembrane domain:**

Amino acids 135-153

**N-glycosylation sites:**

Amino acids 44-48;55-59;64-68

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 171-175

**Tyrosine kinase phosphorylation sites:**

Amino acids 61-69;87-95

**N-myristoylation sites:**

Amino acids 12-18;203-209



550/615

**FIGURE 545**

GGCGGCCCGGAGCTGGGAGCGCGGGGAAGGCGGTTGGGGTTCTGACAGCTGCGCGCGATCCTG  
CTCTCTCTCAGCCGCCTGTGGACATGCGCAAAGGGCCCTCTCCTGAGTCCAG**ATG**ATGCTCAT  
ACCAATGGCTTCAGTGATGGCGGTGACTGAACCGAAATGGGTCTCGGTCTGGAGCCGCTTCCT  
CTGGGTGACGCTGCTGAGCATGGTGCTGGGGTCCCTGCTGGCCCTGCTGCTGCCGCTGGGGGC  
TGTGGAGGAGCAGTGCTTGGCTGTGCTCAAAGGCCTCTACCTGCTCAGGAGCAAACCGGACAG  
GGCGCAGCATGCCGCCACCAAGTGCACCAGCCCGTCCACGGAGCTCAGCATCACCTCCAGGGG  
CGCGACGCTGCTGGTGGCCAAGACCAAGGCCTCTCCAGCGGGTAAGTTGGAAGCCAGAGCTGC  
CCTGAACCAGGCCCTGGAGATGAAGCGCCAGGGCAAGCGGGAAAAAGCCCCAAAAGCTCTTCAT  
GCACGCCCTCAAGATGGACCCGGACTTCGTGGACGCGCTCACCGAGTTTGGCATCTTCTCGGA  
AGAAGACAAGGACATCATCCAGGCGGACTACTTGTACACCAGAGCATTGACCATCTCACCTA  
CCATGAGAAAGCACTGGTCAACCGCGATCGGACACTGCCTCTTGTGGAAGAGATCGACCAGAG  
GTATTTTCAGCATCATCGACAGCAAAGTGAAGAAGGTCATGTCCATCCCCAAGGGGAACTCAGC  
TCTGCGCAGGGTCATGGAGGAGACCTACTACCATCACATCTACCACACAGTGGCCATCGAGGG  
CAACACCCTCACCTCTCGGAAATCAGGCACATCCTGGAGACCCGCTACGCCGTGCCCGGGAA  
GAGCCTGGAGGAGCAGAACGAGGTCATAGGCATGCATGCAGCCATGAAGTACATCAACACGAC  
TCTGGTTTCGCGCATCGGCTCCGTCAACATCAGCGACGTGCTGGAGATCCACAGGCGGGTGCT  
GGGCTACGTGGACCCCGTGGAAGCCGGCAGGTTTCGGACAACACAGGTCCTGGTCGGACACCA  
CATCCCTCCCCATCCGCAGGATGTGGAAAAGCAGATGCAGGAGTTTGTACAGTGGCTCAACTCC  
GAGGAAGCCATGAACCTGCACCCAGTGGAGTTTGCAGCCTTAGCCCATTTATAAACTCGTTTAC  
ATCCACCCTTTTATTGATGGCAACGGGAGGACCTCCCGTCTGCTCATGAACCTCATCCTCATG  
CAGGCGGGCTACCCGCCCATCACCATCCGCAAGGAGCAGCGGTCCGACTACTACCACGTGTTG  
GAAGCTGCCAACGAGGGCGACGTGAGGCCTTTTATTGCTTCATCGCCAAGTGTACTGAGACC  
ACCCTGGACACCCTGCTTTTTGCCACAACCTGAGTACTCGGTGGCACTGCCAGAAGCCCAACCC  
AACCACTCTGGGTTCAAGGAGACGCTTCCTGTGAAGCCCT**TAA**CCCTAGAAATCCTCAGTGACA  
AAGGCTGTCTGAGGTAGGAAA

551/615

**FIGURE 546**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119517
><subunit 1 of 1, 458 aa, 1 stop
><MW: 51778, pI: 7.81, NX(S/T): 2
MMLIPMASVMAVTEPKWVSWSRFLWVTLLSMVLGSLALLLPLGAVEEQCLAVLKGLYL
LRSKPDRAQHAATKCTSPSTELSITSRGATLLVAKTKASPAGKLEARAALNQALEMKRQG
KREKAQKLFMHALKMDPDFVDALTEFGIFSEEDKDIIQADYLYTRALTISPYHEKALVNR
DRTLPLVEEIDQRYFSIIDSKVKVMSIPKGNLSALRRVMEETYYHHIYHTVAIEGNTLTL
SEIRHILETRYAVPGKSLEEQNEVIGMHAAMKYINTTLVSRIGSVTISDVLEIHRRVLGY
VDPVEAGRFRFTTQVLVGHHIPPHQDVEKQMQEFVQWLNSEEAMNLHPVEFAALAHYKLV
YIHPFIDGNGRTSRLLMNLILMQAGYPPITIRKEQQRSDYYHVLEAANECDVRPFIRFIK
CTETTLDTLLFATTEYSVALPEAQPNNHSGFKETLPVKP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-46

**N-glycosylation sites:**

Amino acids 275-279;446-450

**Tyrosine kinase phosphorylation sites:**

Amino acids 216-225;217-225;244-232

**N-myristoylation sites:**

Amino acids 35-41;235-241;266-272;368-374

**Amidation site:**

Amino acids 119-123

552/615

**FIGURE 547**

CCTCTGTCTGTGCTCCCATCCCAGGGAGTATAGGTGGAGCCTCCAGAGCCCATGGACAGGGCA  
TGCTGGGGCTGGGCCAGCCCCAGCGGTGTCTCTAAGGCACCCCTGGGATCCCCACTGAGCTGG  
CCTACTTCAGACAGCCAGGGCCCCACCCCTCTGGCCCCCTTAGTGTCCAGCTCGTGGCCCCCTTG  
GCATTTCCACAAGACGCCAAG**ATG**GAGATTCCCATGGGGACCCAGGGCTGCTTCTCAAAGAGC  
CTCCTGCTCTCAGCCTCAATCCTGGTCCTCTGGATGCTCCAAGGCTCCCAGGCAGCTCTCTAC  
ATCCAGAAGATTCCAGAGCAGCCTCAAAAGAACCAGGACCTTCTCCTGTCAGTCCAGGGTGTC  
CCAGACACCTTCCAGGACTTCAACTGGTACCTGGGGGAGGAGACGTACGGAGGCACGAGGCTA  
TTTACCTACATCCCTGGGATACAACGGCCTCAGAGGGATGGCAGTGCCATGGGACAGCGAGAC  
ATCGTGGGCTTCCCCAATGGTTCCATGCTGCTGCGCCGCGCCAGCCTACAGACAGTGGCACC  
TACCAAGTAGCCATTACCATCAACTCTGAATGGACTATGAAGGCCAAGACTGAGGTCCAGGTA  
GCTGAAAAGAATAAGGAGCTGCCCAGTACACACCTGCCACCAACGCTGGGATCCTGGCGGCC  
ACCATCATTGGATCTCTTGCTGCCGGGGCCCTTCTCATCAGCTGCATTGCCTATCTCCTGGTG  
ACAAGGAAGTGGAGGGGGCCAGAGCCACAGACTGCCTGCTCCGAGGGGGCCAGGGATCTCTGTCC  
ATCTTGCTGCTCGGCTGTATCCCCAGTGCCTTCAGTGACGCCCAGCACATGGATGGCGACCACA  
GAGAAGCCAGAATTGGGCCCTGCTCATGATGCTGGTGACAACAACATCTATGAAGTGATGCCC  
TCTCCAGTCCTCCTGGTGTCCCCCATCAGTGACACAAGGTCCATAAACCCAGCCCGGCCCTG  
CCCACACCCCCACACCTGCAGGCGGAGCCAGAGAACCACCAGTACCAGCAGGACCTGCTAAAC  
CCCGACCTGCCCCCTACTGCCAGCTGGTGCCAACCTTCC**TGAT**GGGTCTGGGCCAGGCCAGC  
CAGGGAGAAGACAAGGCCCCAGCCCTCCTCTGGGAGCCTCACACCTGAGACCAGCAGGACAAG  
GCCATTGGGGGCTGTGGGGCCGATGAGGTGGACTCAGCCAAAGACTCAGCAGCACATGGGGCA  
GGTGTCTTGGCAGGGGGACAGGAGACTGTAACAGGCCCAGGTCTTGTGCAGCCCCCTGAATGC  
ACGCCCCGCTTCGGTCTGTTCCCTCAAGCAAGCTGGCCTGGGCCATGTGCCTGTGAAAGGCAG  
GCTCTGGCCCCCTTTCCATGCCAAAGTCCCCCAAGATCTGGATATCTGGGGACAAGATGGTGGC  
CTCAGGCCTGCCTCCCAGGCAGTTGGCTGGGCTCCCAACTGTCTGTCTCAATGCCCTACCCC  
AACTCCACTAGTGACCCTCAGAGTCTTCTCCCTTAGGACAAGGCAGACACCCCACCATGCGG  
GCCTCAGGTGGCAGAGAGGGCCCAGCCTCACAGGCCTGTGGCCCCACACACCAGTCCCAGCAAG  
GTGACCACGGCTGCTGGACCCCTTCCCTGTTCAAGCAGGCCCAGCCCCCTCTCAGAACCTGCTG  
CCAGCTGCTGGTCTTGGCCCCCACCCTGAATCTTACTGAGTCCCTCTGGGCAGCAGCTCCCTT  
CTCCACCCCACCCCAGCACCCGTCCCAATGTGGCCTCAGCTTGTCTCCCTTCCCCAACT  
ATGCATTTCATTCAGCAATAAATGAGCCTTTGCTGCA

553/615

**FIGURE 548**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119535
><subunit 1 of 1, 300 aa, 1 stop
><MW: 32638, pI: 6.02, NX(S/T): 1
MEIPMGTOGCFKSLLLSASILVLWMLQGSQAALYIQKIPEQPQKNQDLLLSVQGVPDF
QDFNWYLGEETYGGTRLFTYIPGIQRPQRDGSAMGQRDIVGFPNGSMLLRRAPTDSTY
QVAITINSEWTMKAKTEVQVAEKNKELPSTHLPTNAGILAATTIGSLAAGALLISCIAYL
LVTRNWRGQSHRLPAPRGQGSLSILCSAVSPVPSVTPSTWMATTEKPELGPAGDAGDNNI
YEVMPSPVLLVSPISDTRSINPARPLPTPPHLQAEFENHQYQQDLLNPDAPYCVLVPTS
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-32

**Transmembrane domain:**

Amino acids 159-178

**N-glycosylation site:**

Amino acids 104-108

**N-myristoylation sites:**

Amino acids 6-12;29-35;55-61;91-97;157-163;165-171

554/615

**FIGURE 549**

GCCACTCACACCATCTGCTAATGGGACAGCTCACTCTTCCCTCCAAACCATGGCCTTGGCTCA  
AGAGCTTCCTTGTTTCTGGAATGTTCTTTCCTCCAGCTCCAGGTGTTGAAATTCTGCCTGGTC  
TGGGTCTCCTGTTGAAGGACGCCCTCCACTGGGAAGGATCCTCTTGCCCTTACCACCTTGTCTT  
CACCAGCCCCCTGCTCCCCCTCTTCCTCTTGGGGCTGTTGTGCTTGTGATACTTTTTTTTTTGT  
GTGTTTGACACACATCTTCTTCTCACCCCTCTAACACAGTTCTCAACCACAGCACTTTTGTCCC  
TGGAGATGTTGGCAGTGTCCAGAGGCGTGTTGATGGTCCCCTGGGGTGGGGGTGCTGCTGGCA  
CCAGATGGTAGGGAGATGCCAGGGGTGCTGCTCCACACCCTATGGGACACTGCACAGTACACC  
TGGCCTGTGTCCCCACAGCGAGAGCTGGCCCTGGGCAGGCGTGGTCCCTGCGGTGTGTGTTG  
GTTGGGATCCTCCACAGTGACAGACGGTGCGCTCTGCCACGTTTCCACACAGCTCTTTTGCT  
TGTGGAGCTCACCCCTTTCAGAGAGCTCATTTCCCTGCGGTCTTTGGCCTGCAGAAGTAAAA  
**TGA**GGGGTGGTGAATTACACCCCTGCTGGTTACACATGGAAACTCAGGAGTGAGAATTTTGT  
GGAGAGCAAGAGAGGTGAGACTGGGGTGCTGGCTGCCAGCCAGGCGGTCCCTCAGCCCCCTGGA  
GAAGCGGGGTGGGGCCTGCACACCGAGTCCTTCCAGTGAGTCCAGTGATGCTCTCTCCTCTTC  
CTCCAGTCACCTTCTCTCCAGTGCCACTACTGCGCTTTCGATGCAGAATAATTCAGTATTT  
GGCGACTTGAAGTCGGACGAGATGGAGCTGCTCTACTCAGCCTACGGAGATGAGACAGGCGTG  
CAGTGTGCGCTGAGCCTGCAGGAGTTTGTGAAGGATGCTGGGAGCTACAGCAAGAAAGTGGTG  
GACGACCTCCTGGACCAGATCACAGGCGGAGACCACTCTAGGACGCTCTTCCAGCTGAAGCAG  
AGAAGAAATGTTCCCATGAAGCCTCCAGATGAAGCCAAGGTTGGGGACACCCTAGGAGACAGC  
AGCAGCTCTGTTCTGGAGTTCATGTGATGAAGTCCTATCCCGACGTTTCTGTGGATATCTCC  
ATGCTCAGCTCTCTGGGGAAGGTGAAGAAGGAGCTGGACCCTGACGACAGCCATTTGAACTTG  
GATGAGACGACGAAGCTCCTGCAGGACCTGCACGAAGCACAGGCGGAGCGCGGCGGCTCTCGG  
CCGTGCTCCAACCTCAGCTCCCTGTCCAACGCCTCCGAGAGGGACCAGCACCACTGGGAAGC  
CCTTCTCGCCTGAGTGTGCGGGAGCAGCCAGACGTCACCCACGACCCCTATGAGTTTCTTCAG  
TCTCCAGAGCCTGCGGCCTCTGCCAAGACCTAACTCTAGACCACCTTCAGCTCTTTTATTTTA  
TTTTTTTAGTTTTATTTTGCACGTGTAGAGTTTTTGTTCATCAGACAAGGACTTTGATCCTGTC  
CCCTTTGGCATGCGGGAAGCAGCCGCGGGGAGGTAATGAATTGTCTGTGGTATCATGTCAGCA  
GAGTCTCCAAGCCCCACGAACCCTGAGGAGTGGAGTCATACGCGAAGGCCATATGGCCATCGT  
GTCAGCAGAGAGAGTCTCTGTACACAGCCCCGTGAACCCTGAGGAGTGGAGTCATACACGAAG  
GGCGTGTGGCCATCGTGTGAGCAGAGAGAGTCTCTGTACACAGCCCCGTGAACCCTGAGGAGTGG  
AGTCATACGCGAAGGGTGTGTGGCCAGGCTGCAGAGCTGCGTGCCGTTTGTGTCCGAGCATCA  
CGTGTGGCTCCAGCCCTTGTCTGCCAGTGTAGACACCTCTGTCTGCCCCACTGTCTGGGG  
TCGCTCTTGGGAGGCACAGGCATGGGTGTGTCTGGCCTCATTCTGTATCAGTCCAGTGTGTTT  
CTGTATAGTTTGTGTCTCCAGGCAGGCCATGGTAGGGGCCTCGCAGGGGCCATTGGGGAGC  
ACAGGGCCAGGCTGGGGTGAGGAGAGCTCCCCTGTTTTCTGTTTAATTGATGAGCCTGGGAAA  
GGAGTGTGTTCTGCCTGCCCCTTACAGTGGAGCGTTCCGTGTCCATAAAACGTTTTCTAACTG  
GGAA

555/615

**FIGURE 550**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119537
><subunit 1 of 1, 104 aa, 1 stop
><MW: 11136, pI: 8.20, NX(S/T): 0
MLAVSRGVLMVPLGLGVLLAPDGREMPGVLLHTLWDTAQYTWPVSPPTARAGPGQAWSLRC
VLVGILHSDRRRCALPTFPHSSFACGAHPFAESSFPCGLWPAEVK
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**N-myristoylation sites:**

Amino acids 53-59;64-70;97-103

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 74-85

556/615

**FIGURE 551**

CGCCCTTAGC**ATG**GGGTTTCTGCCGCAGCGCCCTGCACCCGCTGTCTCTCTGCTGGTGCAGGCCATCATGCTGGCCAT  
GACCCCTGGCCCTGGGTACCTTGCCCTGCCTTCTTACCCTGTGAGCTCCAGCCCCACGGCTGGTGAAGTGAAGTGA  
GCTGTTCTCTGAAGTCTGTGCCCCACTTCTCCATGGCAGCACCCCGTGGCAATGTCACCAGCCTTTCTTGTCTCTC  
CAACCGCATCCACCACCTCCATGATTCTGACTTTGCCACCTGCCAGCCTGCGGCATCTCAACCTCAAGTGGAA  
GTGCCCCGCGGTTGGCCTCAGCCCCATGCACTTCCCCTGCCACATGACCATCGAGCCCAGCACCTTCTTGGCTGT  
GCCCCACCTGGAAGAGCTAAACCTGAGCTACAACAACATCATGACTGTGCCTGCGCTGCCAAATCCCTCATATC  
CCTGTCCCTCAGCCATACCAACATCCTGATGCTAGACTCTGCCAGCCTCGCCGGCCTGCATGCCCTGCGCTTCT  
ATTCATGGACGGCAACTGTTATTACAAGAACCCTGCAGGCAGGCACCTGGAGGTGGCCCCGGGTGCCCTCCTTGG  
CCTGGGCAGCCTCACCCACCTGTCACTCAAGTACAACAACCTCACTGTGGTGGCCCCGAACCTGCCTTCCAGCCT  
GGAGTATCTGCTGTTGTCTTACAACCGCATCGTCAAACCTGGCGCCTGAGGACCTGGCCAATCTGACCGCCCTGCG  
TGTGCTCGATGTGGGCGGAAATTGCCGCGCTGCGACCAGCCTCCCAACCCCTGCATGGAGTGGCCTCGTCACTT  
CCCCAGCTACATCCCGATACCTTCAGCCACCTGAGCCGTCTTGAAGCCCTGGTGTGAAGGACAGTTCTCTCTC  
CTGGCTGAATGCCAGTTGGTTCCGTGGGCTGGGAAACCTCCGAGTGTGGACCTGAGTGAGAAGTCTCTCTACAA  
ATGCATCACTAAAACCAAGGCCCTCCAGGGCCTAACACAGCTGCGCAAGCTTAACCTGTCTTCAATTACAAAA  
GAGGGTGTCTTTGCCACCTGTCTCTGGCCCTTCTTCCGGAGCCTGGTCCGCCCTGAAGGAGCTGGACATGCA  
CGGCATCTTCTTCCGCTCACTCGATGAGACCAGCTCCGGCCACTGGCCCGCTGCCCATGCTCCAGACTCTGCG  
TCTGCAGATGAAGTTCATCAACCAGGCCAGCTCGGCATCTTCAGGGCCTTCCCTGGCCTGCGCTACGTGGACCT  
GTCCGACAACCGCATCAGCGAGCTTCGGAGCTGACAGCCACCATGGGGGAGGCAGATGGAGGGGAGAAGGTCTG  
GCTGCAGCCTGGGGACCTTGCTCCGGCCCCAGTGGACACTCCAGCTCTGAAGACTTCAGGCCCAACTGCAGCAC  
CCTCAACTTCACCTTGGATCTGTACGGAACAACCTGGTGACCGTGCAGCCGGAGATGTTTGGCCAGCTCTCGCA  
CCTGCAGTGCCTGCGCCTGAGCCACAACCTGCATCTCGCAGGCAGTCAATGGCTCCAGTTCTTCCGCTGACCGG  
TCTGCAGGTGCTAGACCTGTCCACAAATAAGCTGGACCTTACCACGAGCACTCATTACGGAGCTACCACGACT  
GGAGGCCCTGGACCTCAGCTACAACAGCCAGCCCTTTGGCATGCAGGGCGTGGGCCACAACCTTCAGCTTCTGTGGC  
TCACCTGCGCACCCCTGCGCCACCTCAGCCTGGCCACAACAACATCCACAGCCAAGTGTCCAGCAGCTCTGCAG  
TACGTGCTGCGGGCCCTGGACTTCAGCGGCAATGCACTGGGCCATATGTGGGCCGAGGGAGACCTCTATCTGCA  
CTTCTTCCAAGGCCTGAGCGGTTTGATCTGGCTGGACTTGTCCAGAACCGCTGCACACCCTCCTGCCCCAAC  
CCTGCGCAACCTCCCCAAGAGCCTACAGGTGCTGCGTCTCCGTGACAATTACCTGGCCTTCTTTAAGTGGTGGAG  
CCTCCACTTCTGCCCCAACTGGAAGTCTCGACCTGGCAGGAAACAGCTGAAGGCCCTGACCAATGGCAGCCT  
GCCTGCTGGCACCCGGCTCCGGAGGCTGGATGTGAGCTGCAACAGCATCAGCTTCGTGGCCCCCGGCTTCTTTTC  
CAAGGCCAAGGAGCTGCGAGAGCTCAACCTTAGCGCCAACGCCCTCAAGACAGTGGACCACTCCTGGTTTGGGCC  
CCTGGCGAGTGCCCTGCAAATACTAGATGTAAGCGCCAACCTCTGCACTGCGCCTGTGGGGCGGCTTTATGGA  
CTTCTGCTGGAGGTGCAGGCTGCCGTGCCCGGTCTGCCAGCCGGGTGAAGTGTGGCAGTCCGGGCCAGCTCCA  
GGCCCTCAGCATCTTTGCACAGGACCTGCGCCTCTGCTGGATGAGGCCCTCTCCTGGGACTGTTTCCGCCCTCTC  
GCTGCTGGCTGTGGCTCTGGGCCTGGGTGTGCCATGCTGCATCACCTCTGTGGCTGGGACCTCTGGTACTGCTT  
CCACCTGTGCTGGCCTGGCTTCCCTGGCGGGGGCGGCAAGTGGGCGAGATGAGGATGCCCTGCCCTACGATGC  
CTTCGTGGTCTTCGACAAAACGCAGAGCGCAGTGGCAGACTGGGTGTACAACGAGCTTCGGGGGCAGCTGGAGGA  
GTGCCGTGGGCGCTGGGCACCTCCGCCTGTGCTGGAGGAACCGCACTGGCTGCCTGGCAAAACCTCTTTGAGAA  
CCTGTGGGCTCGGTCTATGGCAGCCGCAAGACGCTGTTTGTGCTGGCCACACGGACCGGGTCACTGGTCTCTT  
GCGCGCCAGCTTCTGCTGGCCAGCAGCGCCTGCTGGAGACCGCAAGGACGTCGTGGTGTGGTGTATCCTGAG  
CCCTGACGGCCGCGCTCCCGCTACGTGCGGCTGCGCCAGCGCCTCTGCCGCCAGAGTGTCTCTCTGGCCCCA  
CCAGCCCAGTGGTCAGCGCAGCTTCTGGGCCAGCTGGGCATGGCCCTGACCAGGGACAACCACCACTTCTATAA  
CCGGAACCTCTGCCAGGGACCCACGGCCGA**TAG**CCGTGAGCCGGAATCCTGCACGGTGCCACCTC

557/615

**FIGURE 552**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA119714
><subunit 1 of 1, 1032 aa, 1 stop
><MW: 115799, pI: 8.61, NX(S/T): 12
MGFCRSALHPLSLLVQAIMLAMTLALGTLPAFLPCELQPHGLVNCNWLFKSVPHFSMAA
PRGNVTSLSSLSSNRIHHLHDSDFAHLPSLRHLNLKWNCPVGLSPMHFPCMTIEPSTFL
AVPTLEELNLSYNNIMTVPALPKSLISLSLSTNIMLMLDSASLAGLHALRFLFMDGNCY
KNPCRQALEVAPGALLGLGSLTHLSLKYNLTVVPRNLPSSLEYLLLSYNRIVKLAPEDL
ANLTALRVLDVGGNCRRCDHAPNPCMECPRHFPQLHPDTFSHLSRLEGLVLKDSSLWL
ASWFRGLGNLRVLDLSENFYKCIKTKALQGLTQLRKLNLSFNYQKRVSAHLSLAPSF
GSLVALKELDMHGIFFRSLDETTLRPLARLPMLQTLRLQMNFINQAQLGIFRAFPGLRYV
DLSDNRISGASELTATMGEADGGEKVWLQPGDLAPAPVDTPSSEDFRPNCTLNFTLDLS
RNNLVTVQPEMFAQLSHLQCLRLSHNCISQAVNGSQFLPLTGLQVLDLSHNKLDLYHEHS
FTELPRLEALDLSYNSQPPFGMQGVGHNFSAHLRTRLRHLSLAHNNIHSQVSQQLCSTSL
RALDFSGNALGHMWAEGDLYLHFFQGLSGLIWLDDLSQNRHLHTLLPQTLRNLPKSLQVRL
RDNYLAFFKWWSLHFLPKLEVLDLAGNQLKALTNGSLPAGTRLRRLDVSCNSISFVAPGF
FSKAKELRELNLSANALKTVDSWFGPLASALQILDVSANPLHCACGAAFMDFLLEVQAA
VPGPLPSRVKCGSPGQLQGLSIFAQDLRLCLDEALSWDCFALSLLAVALGLGVPMLHHLCG
WDLWYCFHLCLAWLPWRGRQSGRDEDALPYDAFVVFDKTQSAVADWVYNELRGQLEECRG
RWALRLCLEERDWLPGKTLFENLWASVYGSRKTLFVLAHTDRVSGLLRASFLLAQQRLL
DRKDVVVLVILSPDGRRSRYVRLRQRLCRQSVLLWPHQPSGQRSFWAQLGMALTRDNHFF
YNRNFCQGPTAE
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-30

**Transmembrane domain:**

Amino acids 818-835

**N-glycosylation sites:**Amino acids 64-68;129-133;210-214;242-246;300-304;340-344;  
469-473;474-477;513-517;567-571;694-698;731-735**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 347-351

**Tyrosine kinase phosphorylation site:**

Amino acids 863-871

**N-myristoylation sites:**Amino acids 27-33;41-47;63-69;193-199;361-367;409-415;  
563-569;607-613;695-701;794-800;929-935;945-951;  
1010-1016**Amidation site:**

Amino Acids 974-978

**Leucine zipper patterns:**

Amino acids 204-226;644-666;814-836



558/615

**FIGURE 553**

GGCGTGGGACGTGCTGCGGCGTCCTAGCTGGCTTACAGGGCGGCGGGGGTGTGTGTCTCTCT  
GTTAAGAGTGCTACTCGCCCGGGTTGATCTGTGCATGCCACTCCTGGGTGAGACGGTGAGGT  
CGGCGTCTGCGAGGACGCGGCGGTGGAGTAGAAGGGCAGCCGGAGACAGGCCCGGCGCCCCCTT  
CCGAGGCTAGACGGCCCCAGCTTCGCGGGGATC**ATG**GCATTGCTGGTGGACCGAGTGCGGGGC  
CACTGGCGAATCGCCGCGGGCTCCTGTTCAACCTGCTGGTGTCCATCTGCATTGTGTTCCCTC  
AACAAATGGATTTATGTGTACCACGGCTTCCCCAACATGAGCCTGACCCTGGTGCACCTTCGTG  
GTCACCTGGCTGGGCTTGTATATCTGCCAGAAGCTGGACATCTTTGCCCCAAAAGTCTGCCG  
CCCTCCAGGCTCCTCCTCCTGGCCCTCAGCTTCTGTGGCTTTGTGGTCTTCACTAACCTTTCT  
CTGCAGAACAAACACCATAGGCACCTATCAGCTGGCCAAGGCCATGACCACGCCGGTGATCATA  
GCCATCCAGACCTTCTGCTACCAGAAAACCTTCTCCACCAGAATCCAGCTCACGCTGATTCCCT  
ATAACTTTAGGTGTAATCCTAAATTCTTATTACGATGTGAAGTTTAATTCCTTGGAATGGTG  
TTTGCTGCTCTTGGTGTTTTAGTTACATCCCTTTATCAAGTGTGGGTAGGAGCCAAACAGCAT  
GAATTACAAGTGAACCAATGCAGCTGCTGTACTACCAGGCTCCGATGTCATCTGCCATGTTG  
CTGGTTGCTGTGCCCTTCTTTGAGCCAGTGTGGGAGAAGGAGGAATATTGGTCCCTGGTCA  
GTTTCTGCTTTGCTTATGGTGTGCTATCTGGAGTAATAGCTTTCATGGTGAACCTTATCAATT  
TATTGGATCATTGGGAACACTTCACCTGTCACCTATAACATGTTCCGACACTTCAAGTTCTGC  
ATTACTTTATTCGGAGGATATGTTTTATTTAAGGATCCACTGTCCATTAATCAGGCCCTTGGC  
ATTTTATGTACATTATTTGGCATTCTCGCCTATACCCACTTTAAGCTCAGTGAACAGGAAGGA  
AGTAGGAGTAACTGGCACAACGTCCT**TAA**TTGGGTTTTTGTGGAGAAAAGAATGTTGTCCCA  
AGAAGATAAAAAATATTGTTAAGTGTGCAAGTTATTA

559/615

**FIGURE 554**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA125170
><subunit 1 of 1, 313 aa, 1 stop
><MW: 35066, pI: 9.39, NX(S/T): 5
MALLVDRVRGHWRIAAGLLENLLVSICIVFLNKWIYVYHGFNMSLTLVHFVVTWLGLYI
CQKLDIFAPKSLPPSRLLLLALSFCGFVVFNTLSLQNTTIGTYQLAKAMTTPVIIAIQTF
CYQKTFSTRIQLTLIPITLGVILNSYYDVKFNFGLGMVFAALGVLTSLYQVWVGAKQHEL
QVNSMQLLYYQAPMSSAMLLVAVPFFEPVFEGGGIFGPWSVSALLMVLLSGVIAFMVNLS
IYWIIGNTSPVTYNMFGHFKFCITLFGGYVLFKDPLSINQALGILCTLFGILAYTHFKLS
EQEGSRSKLAQRP
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-27

**Transmembrane domains:**

Amino acids 46-60;75-90;153-167;192-208;221-237

**N-glycosylation sites:**

Amino acids 43-47;92-96;97-101;238-242

**N-myristoylation sites:**

Amino acids 17-23;57-63;140-146;155-161;162-168;283-289

560/615

**FIGURE 555**

GTTAGGCAGAGCCAAGGTGGTTGCAGACCTGGAATCAGAACAGCTTTTAGACCAACCTGAAAG  
CAGGAATGTAAGCACTGTTACAGAGATTTTCGTCTTTGGCTTATTGTGCCTGCAGAGTCTAG  
TGCTTCTTTGCCAGCTGTGCTGACTCAGCACTCC**ATG**CCTGTTTTCTGGAACCAGTCCCTGGA  
GCTGGGCCATGTTTTGATTGACAGTGTGGAGCTAGCCCAGCAAGTACTCTACATGCAACCCCC  
CACCCAGGCACTACCTCTGCTCCTCCTCCATGGCCTCCTGCTACACCGGCAGCTCTATGGAAC  
AAGGCTGCAGGCACACAGGGGGCGCTGGAGTCAAGTGAAGTCTAACCAGGTTCTTCAGACCCA  
AGACCAGCTGTGGGCAAGTCTTAGCAATCCCCGTGCTGCCATGCAAGAGCTGGCTGCTTCAGT  
TTTCTACGGGGTCTCTGCGGGACACTGAGGACAGGGAGGCCCTGATTAGCCTCACACAAGC  
CTGCCTGAGCCCCAGTAGTGGGAGCTGGGTCCAGCCACACACACCTCAGTCTTTGCTGGCCAC  
GCTCATGCCCCCTCCAGCTAAGGGAGCTGGATGCAATGGCAGAGTGCAAGGCCAGATGCACC  
TACTGCCCTCACACCT**TGA**ACCCCGGCTCTGCGGACTGAGTGAGGGCCCCAAGCCTGGCTGT  
TGCGACGCCAGAGTCGCGCTCTCTTGAGTGCCTGTCAGCGGAGTTCACCCGTGTGGGTTCTCTG  
AGTCTCGAAGAGGCGCCAGCTTGCGGAAAGGCGACTGCGGCAACGCCTAGTGCAAGTCAACCG  
GAGGCTGGAGTCACTGCAGGATCTGCTGACCCACGTGATTGCGCAAGACGAGTCCGACGCCCC  
GTGGTCAGTGCTGGGGCCAAATGCACGGCGGCCTCTGGAGGGCGTCTTAGAGACCGAGGCTCT  
AGAAGTGAAGCCAGTTGGTGGGCACGCTACAACGCGACCTTGATTGCCTGTTGCAGCAGCTGAA  
GGGCGCACCCCCGTGCCCCCTCCCGCCGCTGTGCTGCGGTGGCCACGCTCTCTGGACTGGCCG  
CCTACCCCTTGCCCTGGCGACCTCATGCGCCGGCCGGTCCGCAGCCGCCCTGGCACTGGCTGCG  
ACAGTTGTCGCGCCGTGGGCAACTGTTGGTTCGTTACTTGGGCGTGGGCGCGGACGCGAGCAG  
TGATGTACCAGAGCGCGTCTTCCACCTGTCAGCCTTTCGCCACCCGCGCCGCTGCTGCTGGC  
ATTGCGTGGGGAAGCTGCCCTGGACCAGAATGTGCCAGCTCGAATTTCCCTGGTAGCCGAGG  
CTCGGTCTCCAGTCAGCTCCAGTATAAACGTCTGGAGATGAACAGCAACCCTCTGCACTTCAG  
GGTGGAGAATGGTCCAAATCCCACGGTTCCAGAGAGAGGGCTGCTGCTGATCGGGCTACAGGT  
CCTACATGCGGAGTGGGACCCAATAGCTGGAGCCTTGCAAGACAGTCCTTCCAGCCAACCCAG  
CCCTCTGCCTCCCGTCAGCATCAGCACACAGGCCCCGGGCACCAGTGACCTGCCAGCCCCAGC  
CGACCTGACTGTGTACTCGTGTCTGTGTACATGGGAGGGGCCCTTGGCACCGCTAAGCTGCA  
GAGCAGGAACATCGTGATGCATCTGCCTTTACCCACCAAGCTCACCCCCAACACCTGTGTCCA  
AAGGAGGGTCCATGTGTGCAGCCCACCCCTGTCTTGAGCCCGTCTACCAAATAAAGTTGTAG  
TGATTCCA

561/615

**FIGURE 556**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129594
><subunit 1 of 1, 162 aa, 1 stop
><MW: 17598, pI: 6.58, NX(S/T): 1
MPVFWNQSLELGHVLIDSVELAQQVLYMQPPTQALPLLLHGLLLHRQLYGTRLQAHRGR
WSQVTILTQVLQTQDQLWASLSNPRAAMQELAASVFYGGPLGDTEAREALISLTQACLSPS
SGSWVQPHTPQSLLATLMPLPAKGAGCNGRVQGPDAPTALT
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-45

**N-glycosylation site:**

Amino acids 6-10

**N-myristoylation sites:**

Amino acids 97-103;144-150

562/615

**FIGURE 557**

GACCTTGAGCCCTCGAAAGCGAC**ATG**GCGGTTCTCTTAAAGCTGGGCGTTCTCTGCAGTGGCC  
AAGGAGCTCGAGCTCTCCTACTCCGAAGCCGGGTGGTCAGACCCGCTTATGTGTGAGCATTTC  
TCCAGGACCAGCCTACCCAAGGACGGTGTGGTACCCAGCACATTACCTGTCACCAAGCCACC  
ACTCTGGTTCCAAGGCTGCATCTCTCCACTGGACCAGTGAGAGGGTTGTCAGTGTCTGCTCT  
TGGGGCTGATCCCTGCTGGGTACTTGAATCCCTGCTCTGTGGTGGACTACTCTCTGGCTGCAG  
CCCTCACCCCTGCACAGTCACTGGGGCCTTGGACAAGTGGTTACCGACTACGTTTCATGGGGACA  
CCCTGCCGAAGGCTGCCAGGGCAGGCCTCTTGGCACTCTCAGCTTTGACCTTTGCTGGGCTTTGC  
TACTTCAATTACCACGATGTCGGCATCTGCAGAGCGGTTGCCATGCTGTGGAAGCT**CTGA**CCT  
GGGTGCAGCACTTTGATTGTGTGCCTCCTTGCCTCTGCTTTACCAATGCCGTTACCTCGCAG  
TGAGGGGGGATGAAGGATAAGCCCATTGGTGGGCAGAATGTCTTCTAATTACATGGTTATTTT  
CAGAATTTATTTGTTGAGGAAGAGGTTTGAGGAGTTAGGTTTCGACCATTCTGTAGTCTGTGTT  
CCATACTCCACTGAGTGTGGGCAGTAGCTCACAGCCTCGCGGTGAGACTGAACATTTTCATGAG  
CTCATGTTGCCTTTGACCACCATTTCTTAAGGAGAGCCAGCTGATTGCTGTCAGGATAAGAGC  
ATCTCTTCAGCCAGGAGGGAGGCCTGTTCCCTCCTGAGTTAGACTTTGCATGAAGCTCGAAAG  
TATTCCTTTTGGAACCTCCCATTCTTGTTCAGGTGACACCAGCTCTGTTGATGGCTCTGCTTC  
TAGGGAACATTTAATCAGGAGATGCTCTCAATGACTAATTTGTCTAAGTCTTAGGAAGGAGGT  
TGAGGAAAGCTGGATTTAGACAAGTTCAATTTAGGGAGTTCTCCTTGTTTGTGGATTAAAATA  
TGACAGATTGCAAACAGACTACTCTTCAAATGTATCTCAATTGTGCAGAAGTGAGCTGTCCAA  
AAGTATAAGACTAAGTGATAAACTGTCTTCCCACCGTGGGAGTTGTTAATGAGAAAGAAAGTG  
TACTCTGAAAAACAAGGGGG

563/615

**FIGURE 558**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA129793
><subunit 1 of 1, 159 aa, 1 stop
><MW: 17014, pI: 9.38, NX(S/T): 0
MAVLLKLGVLCSGQGARALLRSRVVRPAYVSAFLQDQPTQGRCGTQHIHLSPSHHSGSK
AASLHWTSERVVSLLLLGLIPAGYLNPCSVVDYSLAAALTLHSHWGLGQVVTDYVHGDTL
PKAARAGLLALSALTFAGLCYFNYHDVGICRAVAMLWKL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-15

**Transmembrane domains:**

Amino acids 71-88;126-140

**Glycosaminoglycan attachment site:**

Amino acids 12-16

**N-myristoylation sites**

Amino acids 8-14;58-64;78-84;108-114;148-154

564/615

**FIGURE 559**

CCCAGCCCCGCGTTTCGGCTGCTCTCGAGGAGGCCGGAGTCCCCGGAGACGATGCGCCCCGCGC  
AGCCGCCTGCGCCTGCGGGAGCCGGCTGCCCTTGAGATGGAGTTGCTGCCTCTTTGGCTCTGC  
CTGGGTTTTTCACTTCCTGACCGTGGGCTGGAGGAACAGAAGCGGAACAGCCACAGCAGCCTCC  
CAAGGAGTCTGCAAGTTGGTGGGTGGAGCCGCTGACTGCCGAGGGCAGAGCCTCGCTTCGGTG  
CCCAGCAGCCTCCCGCCCCACGCCCGGATGCTCACCCCTGGATGCCAACCCCTCTCAAGACCCCTG  
TGGAATCACTCCCTCCAGCCTTACCCTCTCCTGGAGAGCCTCAGCCTGCACAGCTGCCACCTG  
GAGCGCATCAGCCGCGCGCCTTCCAGGAGCAAGGTCACCTGCGCAGCCTGGTCCTGGGGGAC  
AACTGCCTCTCAGAGAACTACGAAGAGACGGCAGCCGCCCTCCACGCCCTGCCGGGCCTGCGG  
AGGCTGGACTTGT CAGGAAACGCCCTGACGGAGGACATGGCAGCGCTCATGCTCCAGAACCTC  
TCCTCGCTGCGGTCCGTGTCCCTGGCGGGGAACACCATCATGCGGCTGGACGACTCCGTCTTC  
GAGGGCCTGGAGCGTCTCCGGGAGCTGGATCTGCAGAGGAACATACATCTTCGAGATCGAGGGC  
GGCGCTTTCGACGGCCTGGCTGAGCTGAGGCACCTCAACCTGGCCTTCAACAACCTCCCCTGC  
ATCGTGGACTTCGGGCTCACGCGGCTGCGGGTCCTCAACGTCAGCTACAACGTCCTGGAGTGG  
TTCCTCGCGACCGGGGGAGAGGCTGCCTTCGAGCTGGAGACGCTGGACCTGTCTCACAACCAG  
CTGCTGTTCTTCCCGCTGCTGCCCCAGTACAGCAAGTTGCGGACCCTCCTGCTGCGCGACAAC  
AACATGGGCTTCTACCGGGACCTGTACAACACCTCGTCGCCGAGGGAGATGGTGGCCAGTTC  
CTCCTCGTGGACGGCAACGTGACCAACATCACCACCGTCAGCCTCTGGGAAGAATTCTCCTCC  
AGCGACCTCGCAGATCTCCGCTTCCTGGACATGAGCCAGAACCAGTTCAGTACCTGCCAGAC  
GGCTTCCTGAGGAAAATGCCTTCCCTCTCCACCTGAACCTCCACCAGAATTGCCTGATGACG  
CTTCACATTCGGGAGCACGAGCCCCCGGAGCGCTCACCGAGCTGGACCTGAGCCACAACCAG  
CTGTGCGAGCTGCACCTGGCTCCGGGGCTGGCCAGCTGCCTGGGCAGCCTGCGCTTGTTCAAC  
CTGAGCTCCAACCAGCTCCTGGGCGTCCCCCTGGCCTCTTCGCCAATGCTAGGAACATCACTAC  
ACTTGACATGAGCCACAATCAGATCTCACTTTGTCCCTGCCAGCTGCCTCGGACCGGGTGGG  
CCCCCTAGCTGTGTGGATTT CAGGAATATGGCATCTTTAAGGAGCCTGTCTCTGGAGGGCTG  
TGGCCTGGGGGCATTGCCAGACTGCCCATTCCAAGGGACCTCCCTGACCTACTTAGACCTCTC  
AAGCAACTGGGGGGTTCTGAATGGGAGCCTCGCCCCACTCCAGGATGTTGCCCCCATGTTACA  
GGTCCTGTCTCTCAGGAACATGGGCCTCCACTCCAGCTTTATGGCGTTGGACTTCTCTGGGTT  
TGGGAATCTCAGGGACTTAGATCTGTGCGGGGAATTGCTTGACCACCTTCCAAGGTTTGGGGG  
CAGCCTGGCCCTGGAGACCCTGGATCTCCGTAGAAACTCGCTCACAGCCCTTCCCCAGAAGGC  
TGTGTCTGAGCAGCTCTCGAGAGGTCTGCGGACCATCTACCTCAGTCAGAATCCATATGACTG  
CTGTGGGGTGGATGGCTGGGGGGCCCTGCAGCATGGGCAGACGGTGGCCGACTGGGCCATGGT  
CACCTGCAACCTCTCCTCCAAGATCATCCGCGTGACGGAGCTGCCGGAGGTGTGCCTCGGGA  
CTGCAAGTGGGAGCGGCTGGACCTGGGCCTGCTCTACCTCGTGCTCATCCTCCCCAGCTGCCT  
CACCTGCTGGTGGCCTGCACTGTCATCGTCCTCACTTTTAAGAAGCCTCTGCTTCAGGTCAT  
CAAGAGCCGCTGCCACTGGTCCCTCCGTTTACTGACCTGGCTGTGTGCCAAGACTCGAAATTCTG  
GTCCGCACACAACAGGACACTTTCTCTGCCAGCTTTCAAGATGTGATGCAGAGGCCAAGTCTG  
ACGAATTGAAGTTTCAATTAAAATTTAATATGTTTCCATTCCCTCATCGCCACCCCAACCCCG  
CCCCACCACCGCCCAAGTTCTTTTCCATCATTATAATTATCCTTATTATCTTGGTAAAT  
ATTTATTAAGTGACTTTTTTCAGAAATAAAAGGCAACGTGTCTCATAAATATTTTTTAAAAAA  
AAAAAAAAAAAAA

565/615

**FIGURE 560**

><subunit 1 of 1, 692 aa, 1 stop  
><MW: 76366, pI: 6.07, NX(S/T): 11  
MELLPLWLCCLGFHFLTVGWRNRSGTATAASQGVCKLVGGAADCRGQSLASVPSSLPPHAR  
MLTLDANPLKTLWNHSLQPYPLLESLSLHSCHLERISRGAFQEQGHLRSLVLGDNCLSEN  
YEETAAALHALPGLRRLDLSGNALTEDMAALMLQNLSSLRSVSLAGNTIMRLDDSVFEGL  
ERLRELDLQARNYIFEIEGGAFDGLAELRHLNLAFFNNLPCIVDFGLTRLRLVNLVSYNVLEW  
FLATGGEAAFELETLDLSHNQLLFFPLLPQYSKLRITLLLRDNNMGFYRDLYNTSSPREMV  
AQFLLVDGNVTNITTVSLWEEFSSSDLAFLDLMSQNFQYLPDGFRLKMPSLSHLNLH  
QNCLMTLHIREHEPPGALTELDLSHNQLSELHLAPGLASCLGSLRLFNLSNQLLGVPFG  
LFANARNITTLDMSHNQISLCPLPAASDRVGPSPCVDFRNMASLRSLSLEGCGLGALPDC  
PFQGTSLTYLDLSSNWGVNLGSLAPLQDVAPMLQVLSLRNMGLHSSFMALDFSGFGNLRD  
LDLSGNCLTTFPRFGGSLALETDLRRNSLTALPQKAVSEQLSRGLRTIYLSQNPYDCCG  
VDGWGALQHGQTVADWAMVTCNLSSKIIRVTELPGGVPRDCKWERLDLGLLYLVLIPLSC  
LTLVACTVIVLTFKKPLLQVIKSRCHWSSVY

**Important features of the protein:****Signal peptide:**

Amino acids 1-18

**Transmembrane domain:**

Amino acids 651-672

**N-glycosylation sites:**Amino acids 21-25;74-78;155-159;232-236;292-296;309-313;  
312-316;408-414;427-431;500-504;622-626**Glycosaminoglycan attachment site:**

Amino acids 533-537

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 566-570

**N-myristoylation sites:**Amino acids 24-30;39-45;45-51;141-147;199-205;245-251;  
308-314;396-402;416-422;420-426;471-477;  
484-490;497-503;522-528;545-551;555-561;610-616**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 657-668

**Leucine zipper patterns:**

Amino acids 48-70;492-514



566/615

**FIGURE 561**

TGGCCTACTGGAAAAAAAAAAAAAAAAAAAAAGTCACCCGGGCCCCGCGGTGGCCACAACAT  
GGCTGCGGCGCCGGGGCTGCTCTTCTGGCTGTTTCGTGCTGGGGGCGCTCTGGTGGGTCCCGGG  
CCAGTCGGATCTCAGCCACGGACGGCGTTTCTCGGACCTCAAAGTGTGCGGGGACGAAGAGTG  
CAGCATGTTAATGTACCGTGGGAAAGCTCTTGAAGACTTCACGGGCCCTGATTGTCGTTTTGT  
GAATTTTAAAAAAGGTGACGATGTATATGTCTACTACAACTGGCAGGGGGATCCCTTGAAC  
TTGGGCTGGAAGTGTTGAACACAGTTTTGGATATTTTCCAAAAGATTTGATCAAGGTA  
TAAATACACGGAAGAAGAGCTACATATTCAGCAGATGAGACAGACTTTGTCTGCTTTGAAG  
AGGAAGAGATGATTTTAAATAGTTATAATGTAGAAGAGCTTTTAGGATCTTTGGA  
CTCTGTACCTGAAGAGTCGAAGAAAGCTGAAGAAGTTTCTCAGCACAGAGAGAAATCTCCTGA  
GGAGTCTCGGGGGCGTGAACCTGACCCTGTGCCTGAGCCCGAGGCATTCAGAGCTGATTCAGA  
GGATGGAGAAGGTGCTTTCTCAGAGAGCACCGAGGGGCTGCAGGGACAGCCCTCAGCTCAGGA  
GAGCCACCCTCACACCAGCGGTCCTGCGGCTAACGCTCAGGGAGTGCAGTCTTCGTTGGACAC  
TTTTGAAGAAATTCTGCACGATAAATTGAAAGTGCCGGAAGCGAAAGCAGAACTGGCAATAG  
TTCTCCTGCCTCGGTGGAGCGGGAGAAGACAGATGCTTACAAAGTCCTGAAAACAGAAATGAG  
TCAGAGAGGAAGTGGACAGTGCCTTATTACAGCAAAGGATTTTCGTTGGCATCAAAATCT  
AAGTTTGTTTTACAAAGATTGTTTTTTAGTACTAAGCTGCCTTGGCAGTTTGCATTTTTGAGCC  
AAACAAAAATATATTATTTTCCCTTCTAAGTAAAAAAAAAAAAAAAAAAAAA

567/615

**FIGURE 562**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131639
><subunit 1 of 1, 303 aa, 1 stop
><MW: 33900, pI: 4.81, NX(S/T): 2
MAAAPGLLEWLFVLGALWWVPGQSDLSHGRRFSDLKVCGDEEC SMLMYRGKALEDFTGPD
CRFVNFKKGDDVYVYYKLAGGSLELWAGSVEHSFGYFPKDLIKVLHKYTEEELHIPADET
DFVCFEGGRDDFNSYNVEELLGSLELEDSVPEESKKAEEVSQHREKSPPEESRGRELDVPV
EPEAFRAEDGEDGAFSESTEGLQGQPSAQESH PHTSGPAANAQGVQSSLDTFEEILHDK
LKVPGSESRTGNSSPASVEREKTDAYKVLKTEMSQ RSGQCVIHYSKGFRWHQNLSLFYK
DCF
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-22

**N-glycosylation site:**

Amino acids 294-298

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 30-34

**Tyrosine kinase phosphorylation site:**

Amino acids 67-76

**N-myristoylation sites:**

Amino acids 205-211;225-231;277-283

**Amidation site:**

Amino acids 28-32

568/615

**FIGURE 563**

GCCAGCCGTGGGATTAGGCTTCGCCGGCTACGATTGCGGCCCCCATCTTCTGACTTTTCCTCG  
TGTGACCCATCTTTTCAAATTCCTTACCTGAGGAAGGAGCCCGATTACAAGGATATTTACCT  
GCTCCTACCTGATCTAGGGACGAGGATGGGAAGACCGCCTGTGGCCATGAGCCCTCCCCGGT  
GCTCCTGGGGCTAAGGCTGGGGCTGCAGCC**ATG**GGGCTGGGTCAGCCCCAGGCCTGGTTGCTG  
GGTCTGCCACAGCTGTGGTCTATGGCTCCCTGGCTCTCTTCACCACCATCCTGCACAATGTC  
TTCCTGCTCTACTATGTGGACACCTTTGTCTCAGTGTACAAGATCAACAAAATGGCCTTCTGG  
GTCGGAGAGACAGTGTTTCTCCTCTGGAACAGCCTCAATGACCCTCTCTTCGGTTGGCTCAGT  
GACCGGCAGTTCCTCAGCTCCCAGCCCCGCCTGTGTGGAGAGGAGCTGCTTGTGGGCAGTGAG  
GAGGCGGACAGCATCACCTTGGGCCGGTATCTCCGGCAGCTGGCACGCCATCGGAACCTCCTG  
TGGTTCGTGAGCATGGACCTGGTGCAGGTGCAGTGGCTCACGCCTGTAATCCCAGCACTTCGG  
GACGCCAAGGTGGAAAGACCGCTTGAGCCCAGGAGTTCGAGGCTGCAAT**TGA**GTTATGATTGCA  
CCACTGCACTCCAGCCTGGGCGGCAGAGAAAGGCTCCATCTCTAAAAAAGAAGAGCTAAGTG  
CTGTACCTAAAACATGCAGTATATAAACTGGCTGAACTTAGAAATAAACTGTTTTCATGTTAT  
GAAAA

569/615

**FIGURE 564**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131649
><subunit 1 of 1, 153 aa, 1 stop
><MW: 17603, pI: 7.42, NX(S/T): 0
MGLGQPQAWLLGLPTAVVYGSLALFTTILHNVFLYYVDTFVSVYKINKMAFWVGETVFL
LWNSLNDPLFGWLSDRQFLSSQPRLCGEELLVGSEEADSITLGRYLRQLARHRNFLWFVS
MDLVQVQWLTPVIPALRDAKVERPLEPRSSRLQ
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**N-myristoylation sites:**

Amino acids 4-10;12-18;93-99

**Leucine zipper pattern:**

Amino acids 102-124

570/615

**FIGURE 565**

CGGCACGAGTAAAATGGAGATAATATCACCATGCACTCAGCCCTAGCCACTGCATTGCTGTTA  
CTGATACCATTACTGCTGCTACGTCGTTTTTTTGGATGGCTCAGCCCTTAGGGAAGGGGGATCA  
AGGGAGAAGCCCGGACCTTCCCGCAGGAGGTGGGCTGGGCACAGCCCTGAACCATGGAGGTCA  
CCCACCCTGAGGTCGGGACCTGGGTTCCCTTCCTATCCACTGGGGGTCCCAGCCTTTGTCTTC  
ATCTCTCCAGGTCCCAGCCCTTCACAGTGGGCACTTCCCTGCCTGTGACGGAGGCCCCAGCCA  
TCTCC

571/615

**FIGURE 566**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131652
><subunit 1 of 1, 89 aa, 1 stop
><MW: 9688, pI: 11.49, NX(S/T): 0
MHSALATALLLLLIPLLLLRFFDGSALREGGSREKPGPSRRRWAGHSPEPWRSPTLRSGP
GFPSYPLGVPAFVFISPGSPSQWALPCL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-18

**Glycosaminoglycan attachment site:**

Amino acids 58-62

572/615

**FIGURE 567**

AGTCTAGCAGGAAAGGAGAGGGAGCTTTCCCCGAAGACCCTCCTGGACCAGCCCCAGGCTCCT  
GTGCTGGTTGCACGCCAGGGCCTGTACTGACCACCTCCACGTGCCACTGGGGCTGTAAGGAGGA  
**ATG**GCGGCCGTGGGCAGCCTGCTTGGCCTGGCAGCCTCTTCCTGGCTAGGGGGCCAGAACGCC  
TCTGACCACAGCCTGTGGCTCCTGAGGAAGCCCCGAGGCTCATCCTGCCCCGGCACGGGTAC  
CAGCTCTGCCGGCTGAGGCAGAGCACCGTGAAGGCCACCGGACCTGCACTCCGCCGCTGCAC  
ACATCCTCCTGGCGAGCTGACAGCAGCAGGGCCTCACTCACTCGTGTGCACCGCCAGGCTTAT  
GCACGACTCTACCCCGTGCTGCTGGTGAAGCAGGATGGCTCCACCATCCACATCCGCTACAGG  
GAGCCACGGCGCATGCTGGCGATGCCATAGATCTGGACACCCTGTCTCCTGAGGAGCGCCGG  
GCCAGGCTGCGGAAGCGTGAGGCTCAGCTCCAGTCGAGGAAGGAGTACGAGCAGGAGCTCAGT  
GATGACTTGCACTGTGGAGCGCTACCGACAGTTCTGGACCAGGACCAAGAAG**TGA**CCGTGGCTC  
CAGCCACCCCGGGACATTGCTAAGATGGGAGGGCTGTTCTTAAATCACTCGTTCTTGAAGCTGC

573/615

**FIGURE 568**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA131658
><subunit 1 of 1, 164 aa, 1 stop
><MW: 18903, pI: 11.08, NX(S/T): 1
MAAVGSLLGLAASSWLGGQNASDHSWLLRKPRGSSCPGTGHQLCRLRQSTVKATGPALR
RLHTSSWRADSSRASLTRVHRQAYARLYPVLLVKQDGSTIHIRYREPRMLAMPIDLDTL
SPEERRARLRKREAQLQSRKEYEQELSDDLHVERYRQFWTRTKK
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-18

**N-glycosylation site:**

Amino acids 20-24

**N-myristoylation sites:**

Amino acids 5-11;9-15;17-23;18-24



574/615

**FIGURE 569**

GGTGCCAAGGGTTCGGGGGGGAGCACTGAGGCTTTAGCAGCTCTCCTGTATCCTCATTTGCAT  
CCTCCTGTAGCAGCTGGAAAATTCAGATTACAGGTGAAATTCCTGGCTGGCAATCTTCTGTA  
TATGGACACAGTGATGTGCCAGAAGGGCTTTGCATCCCTGAGACTGAAGGAAGCTCCATTTTT  
GGAGCCCTCCCACACCTTGCTCTGTGTGCCTCTCATTCTGATTTGAATTCCTATTTTGCTATA  
TGATGAAGCTGTAATCCTAAGTTTAAAAAGGGGAGTAGGTATTGACATCATGGTAGAAATAGG  
CTGTCTTATGGAAGTGTAGTTAGGGATCACAGCCTATTGGACCAGCCCCAGCCTTAGCAGCAG  
TTCTGTACACTGATTCTTCCAGATTAGTCTACGTTCCCTCGAACAGACCTATGCCATGGGTTA  
CAACTACAATTTGTTGTCGATTAGAGTTAACTTACAGACTCTCAAAACCCCATTTCTTTGGGTT  
TAGGCAACTTCCAGAAGTAGTCATTTATTTGAATTTTAGTCTAAGATCAACTGAATTAGGGAG  
GTTTGAAAGTGTAAGCAAAATCGTACATTCCTCAAAACACTTTGTAAAGAAGGAATGGGTAGTG  
TCAACTAAAGGAAATGGTGTGCATCCCAGCAAAAGAAAGAGACCGAAAGCAAAGTCATAAACC  
**ATG**CCCACGAGCTCAGCTGTCTGCTCCGTGTCTCTCCATACCCTTGTTGACTGTGCTCATA  
TTAGCCAGAGACCTAAGTGCTCTTGGAGGATGTCCCTGGGGCCCCCTCCCCCTCCGCTGTAC  
TGTCTACTTCCCTGATCCTCTCTTCTGTGCAGGAGAGGTCCAGGCCTTCTATGAGGACCTGAGT  
GGCCGGCAGTACGTGAATGAAGTCTTCAACTTCAGCGTGGACAAGCTCTATGACCTCCTCTTC  
ACCAACTCGCCCTTCCAGCGGGATTTTCATGGAGCAGCGGCGCTTCTCTGATATCATCTTCCAT  
CCATGGAAAAAGGAGGAGAATGGAAACCAGAGCCGAGTGATTCTTTACACCATCACCTTACC  
AACCCTCTGGCTCCCAAACTGCCACTGTCAAGGAGACACAGACCATGTACAAGGCGAGCCAG  
GAGAGTGAATGTTACGTGATAGATGCCGAAGTCTCACCCACGACGTGCCCTACCACGACTAC  
TTCTACACAATCAATCGCTACACGCTCACCCGTGTGGCTCGGAACAAGAGCCGACTCAGGGTC  
TCCACAGAGCTGCGCTATCGAAAACAGCCCTGGGGGTTAGTGAAAACGTTTCATCGAGAAGAAC  
TTCTGGAGTGGGCTGGAGGACTACTTCCGCCATTTAGAGAGCGAGCTGGCCAAAACGAGAGC  
ACTTATTTGGCTGAGATGCACAGACAATCTCCCAAAGAGAAGGCCAGCAAGACTACAACGGTG  
CGGAGGAGGAAGCGTCCCCATGCCACCTGCGAGTCCCTCACCTGGAAGAGGTGATGAGCCCG  
GTCACCACGCCCACAGATGAGGATGTGGGCCACAGGATCAAACATGTGGCAGGTTCACACAG  
ACGCGGCATATCCCGGAGGACACCCCCAACGGTTTCCACCTGCAGAGCGTGTCCAAGCTGCTG  
CTGGTTATCAGCTGTGTTCTGGTGCTGCTGGTCATCCTTAACATGATGCTCTTCTACAAACTC  
TGGATGTTGGAATACACCACGCAGACCCCTCACTGCCTGGCAGGGTCTAAGGCTCCAAGAAAG  
TTACCCAGTCTCAGACAGAATGGGGCCAGCTCTTAGAGTCCCAACAAAAGTACCACGATACT  
GAGCTCCAAAATGGAGGGAAATCATCAAATCCTCAGTGATGCTCCTTGACCAGATGAAGGAC  
TCGCTCATCAACCTTCAGAACGGCATCAGGTCCCGCGACTACACGTCGGAAAGTGAAGAAAAG  
AGGAATCGCTATCAT**TGA**CAAGGCAGGAACAGGGTGGCTGCAAGAGGCCTGTGCAATACATGT  
ACATAGACCATATAAATATATATATAAATATATATATATACAGAATATAAATATATATATT  
ATATACAGATTTTAAAAAAGAGATAATGCCTATGTACCAGGGAGAAGGAGCGGGCCCTCCCGC  
GCCCTGTGCTGGCCGGAGCAGCGTTTTCTTATGGTGGAGCAGCTGAGGAGGGCAGGAACCGCC  
TCTCAGCACCGACCTCCCCTGATCTCCCTCCTCCACCCCTCTGTTCCCCACCCCTTCCCTTGC  
TGGCCATTCTTGGCTTTTAGAAGGGAAATGTTGAGCCAAAGTTATGCCTGCGAAGACCCTAAG  
GTCTCAAAAAGAAGTCTTAAGACGGCATTGCTTAAGGTGCTTCATTCCCTAATCCCTTTTGA  
TTTGTTCCTCAAAATAAAAGAGAATCTTTTCTTCCCTAAAAAAAAAAAAAAAAAAAAAAAAA

575/615

**FIGURE 570**

><subunit 1 of 1, 425 aa, 1 stop  
><MW: 49786, pI: 8.84, NX(S/T): 3  
MPTSSAVLLRLVLSIPLLTVLILARDLSALGGCPWGPLPLRCHCLLPDPLFCAGEVQAFYE  
DLSGRQYVNEVFNFSDKLYDLLFTNSPFQORDFMEQRRFSDIIFHPWKKEENGNSQSRVIL  
YTITLTNPLAPKTATVRETQTMKASQSECEYVIDAEVLTHDVPYHDYFYTINRYTLTRV  
ARNKSRLRVSTELRYRKQPWGLVKTFFIEKNFWSGLEDFRHLESELAKESTYLAEMHRQ  
SPKEKASKTTTVRRRKRPHAHLRVPHLEEVMSPTTPTDEDVGHRIKHVAGSTQTRHIPE  
DTPNGFHLQSVSKLLLVISCVLVLLVILNMMLFYKLWMLEYTTQTLTAWQGLRLQERLPQ  
SQTEWAQLLESQOKYHDTLQKWREIIKSSVMLLDQMKDSLINLQNGIRSRDYTSESEEEK  
RNRYH

**Important features of the protein:****Signal peptide:**

Amino acids 1-28

**Transmembrane domain:**

Amino acids 312-334

**N-glycosylation sites:**

Amino acids 73-77;114-118;183-187

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 97-101

**Tyrosine kinase phosphorylation sites:**

Amino acids 144-153;188-196

**N-myristoylation sites:**

Amino acids 201-207;291-297

**Leucine zipper pattern:**

325-347

576/615

**FIGURE 571**

GTAGAGAGTGAAGCAGCAAGACTGCAGAGCCTCATCAAGAAGTGTGGAGTGAAGGGAAGGCTTCAGATGGACAAT  
TTGTGTGCTGGGGAAAAAATGGAATGTGCTGCAAATTCCTTGGGATAAGGGTGGACGGCTGCTCTGTCAACTT  
TGACCATTTTCAGATTCTGCGGGCCATTGGTAAAGGGAGTTTGGAAAGGTATGCATCGTGAGAAGCGAGACAC  
TAAGAAAATGTATGCAATGAAGTACATGAACAAGCAGAAGTGCATCGAGAGGGATGAGGTTCCGAATGTTTTCCG  
GGAGCTGCAGATCATGCAAGGGCTGGAGCACCCCTTCTGGTCAATCTGTGGTACTCCTTCCAGGATGAGGAGGA  
CATGTTTCATGGTGGTGGACCTGCTCCTGGGAGGCGACCTGCGCTACCATCTGCAGCAGAATGTGCATTTACAGA  
GGGGACTGTGAACTCTACATCTGTGAGCTGGCACTGGCCCTGGAGTATCTTCAGAGGTACCACATCATCCACAG  
AGACATCAAGCCAGACAATATCCTGCTGGATGAACACGGACATGTTACATTACAGACTTCAACATAGCGACGGT  
AGTGAAAGGAGCAGAAAGGGCTTCTCCATGGCTGGCACCAAGCCCTACATGGCTCCAGAAGTATTCAGGTGTA  
CATGGACAGAGGCCCCGGATACTCGTACCCTGTGCACTGGTGGTCCCTGGGCATCACAGCCTATGAGCTGCTGCG  
GGGCTGGAGGCCGTACGAAATCCACTCGGTACGCCCATCGATGAAATCCTTAACATGTTCAAGGTGGAGCGTGT  
CCACTACTCCTCCACGTGGTGCAGGGGATGGTGGCCCTGCTGAGGAAGCTCCTGACCAAGGATCCTGAGAGCCG  
CGTGTCCAGCCTTCATGACATACAGAGCGTGCCCTACTTGGCCGACATGAACTGGGACGCGGTGTTCAAGAAGGC  
ACTGATGCCCGGCTTTGTGCCAATAAAGGGAGGTTGAACTGCGATCCACATTTGAGCTTGAAGAGATGATTCT  
AGAATCCAAGCCACTTCACAAAAGAGAAGCGATTGGCAAAGAACAGATCCAGGGATGGCACAAGGACAGCTG  
CCCCTGAATGGACACCTGCAGCACTGTTTGGAGACTGTCCGGGAGGAATTCATCATATTCAACAGAGAGAAGCT  
CAGGAGGCAGCAGGGACAGGGCAGCCAGCTCTTGGACACCGACAGCCGAGGGGAGGCCAGGCCAAAGCAAGCT  
CCAGGACGGGTGCAACAACAACCTCTCACCCACACCTGCACCCGTGGCTGCAGCAGCTGAGGCCACACTTGTG  
CTGCTCAACAGGACTGCACTCGTCTCTGCCCTGCCACCCAGAGCCCCTCTTTGTGCCCTGATGGTCCCTGTCTC  
ACCCCTGAAAACATCAGATGCAGAAAAAGCCCTGGACTTGGAGCTGGGAAGCCTGGGTTCTGGTCCCATCTCCAT  
GACTGATTCACGTGTGACCTCAGACAAGTCACGCCCTCTCTGTGCCCTCCGTTTTCTGCATCTGCCAAAGGGTTA  
AACACTTCTGCCCCACTTCAAATTACAAGATTATGGGGAGAACCCTAATTAGGTAGGAAACATGAAAACCTTTGA  
TATTTATAAAATCATTTTACGTGCAAAATATAACCTTAATATTTGAAGTGACCCCCATTCCCCAAAGCAATCAA  
ACCGTCATGACTTTGCAATTTGGCACATCCTAGCTTGTAGAGGGCACTTCCGAAAAACACAGCCCTGACAGCAA  
AATAAAGGTCTGATATGTTGGCCCTTCTATGGAACAACGCTGCCAAATCCTGGAGCAAAACCTGAAGTGTCTT  
CATGTGCATTCTCTGGCAGGCCACAGTCCTTCTGAGCTTGAAGATGGTGCAGCATGCAGACCAGACTTGTCCCC  
AAGGTCTCAGCGCTGCGGTCTCACTCCTCCCTCATTTAAGAAGACTATCCTTACCTTTTAGTTTCAGCAGTCCT  
CACCACCACCATATCCCCAGTGCTGGGATGGCACACAGGTGTCCATTAGATGAGAGTTGGGTGCTGAGCATTG  
GTTACTCCTGCAGAGTGTAATCAGCACCCCATCCAAGTGGCCCGAAAGCCAGACCTGCAGCAGAACTCTCCAAC  
TCTCTATCAGCTTTTCAAGGTTTCTCTCCTGGGAAGGGTGAAAATCAGCTTGTGAGATTCTTCTTACAGAGAGT  
ATCCAATCGGTATTGGTGGAGCGGCTCCCTATTTATACAATAGGAAGCATGGGTGCTTAGAAAGTTTATTTCAGG  
AGGAAAATGGGTTACACAAAAAGCAAACCTACATTCTGATCTGCTCAGGGAGAAGCTTGCCTTTGAACTGGAAGA  
TGTTGGGATGAGCAGGGAAGCTTAGACTTTGGAGTCAGGTTTGTGTTTCAAGATCCAGCCCTGCTGGCTACTAAC  
TAACTGGGAGACCTTAGGCAAGCATGCAATCGCTCTGAATGGCAGTTTCTCATTTTTTAAACAGGGATAATAAA  
ACTAATATTGCAGGGGAGTTACAGGGTTAAATAAGATCCTGTGTGTAACCCCAAGCATTGGATGACTCATAGAAT  
GGCCTTTTTTGTGAGCATAATCGTCATCATTATTTAGATACTTTCTCCTTCACTCACCCAGCAGGTGAGTTTTT  
TGTGCAAAACAAACCTGTTTAGGATTCTTCCAAATGTTCTTCTGGGGTCTTTGATATTTGTTTGTACATCCTGC  
TGAAGTTCGACTGTGTTTTTATTTTTTATCCAACTTCCATTTTTTCACTTTTTTACATGATTACTCAATCCTTGGG  
GCTGTCCATGTCTCTTAGATTTCTTAAAGACATTTTAAATGTATGGTTAGGTTTTATATTTTTATTTTTTAA  
AAAAGAAATAGTCAGTGTTTTCTCCTTTCAACCGAGACTATTTCTGGATTGTGTGCTCCTCGTCAGTTGACTTGT  
TTTGCACACTTTTCTTACTTCATGTCCCATCAACAACCGTCTGCTCCCCACCTCCCCAGGAAATAAGGGGC  
CTGCTCCTCTCCCTACTGTGACCCCTGGAGGCTCTTAAGATGATGATGGTTTTTTTTATTGGGCTGAGTTCACGAA  
TTAGGGGCAGGAGCTGGAAGTCGCCCTAGGAACACCAGATTTCTGGTTCTGTTCAAGTTGGCATTCTTGTGTTG  
GAATAAACTATTTCTTG

577/615

**FIGURE 572**

```
</usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA136110
<subunit 1 of 1, 364 aa, 1 stop
<MW: 42195, pI: 7.40, NX(S/T): 1
MKYMNKQKCIERDEVNRNVFRELQIMQGLEHPFLVNLWYSFQDEEDMFMVVDLLLGGDLRY
HLQQNVHFTEGTVKLYICELALALEYLQRYHIIHRDIKPDNILLDEHGHVHITDFNIATV
VKGAERASSMAGTKPYMAPEVFQVYMDRGPYSGSYPDWWSLGITAYELLRGWRPYEIHVS
TPIDEILNMFKVERVHYSSTWCKGMVALLRKLLTKDPESRVSSLHDIQSVPYLADMNWDA
VFKKALMPGFVPNKGRLNCDPTFELEEMILESKPLHKKKKRLAKNRSRDGTDSCPLNGH
LQHCLETVREEFIIIFNREKLRRQQGQGSQLLDTSRGGGQAQSKLQDGCNNNLLTHTCTR
GCSS
```

**Important features of the protein:****N-glycosylation site:**

Amino acids 285-289

**N-myristoylation sites:**

Amino acids 123-129;290-296;337-343;339-345;348-354

**Serine/Threonine protein kinases active-site signature:**

Amino acids 92-105

**FIGURE 573**

CTCCAGTTTCGCCGACTGTAACATGTTTTCATCCAGTTCAGTATGTTTTGTATGCAAGTTGGAAATAAATAAACGCTC  
CTGAACTGGATGAAACATGTTACAGTCGCCCGAAACATGAGAGGCTGTGTGAGAAGCTGCAGCCGCCGCAGAGG  
AGACCTCAGCATCATCTAGAGCCCAGCGCTGGCCCTGCCTCCGCCTGCGCCGCCGCCGCGCTCGCCGTTTTCTGTT  
CCTGCTACTGTCCCACCTAAACAACATCCCGTTACACGGACAAGTGAACATCTGTGGCTGTCCCTCTCCTTTCTTC  
CTCCTCTTCCAACCTCCTTCTCCTCCTCCCACTCCCGCAGCAGCAAGAGCCCCCACTGACGCTGGCA  
CAACTGCAAAACGGGTGCTATCCGCACAACATTTATCTCGTCTCCTCGGCTCCCCAAGGCATTGGACCCATCGCCGC  
GTCTTTTATTTTTGCAAAGTTGCATCGCTGTACATATTTTTGTCCCGCCACCTCCTCTGTCTCTGGAGTGCCC  
TACAGCCCCGCAAACCTCCTCCTGGAGCTGCGCCCTAGTGCCCTGTCTGGGCAGTGGCGTTCCCCCCCATCCTCCC  
GCGCCAGCCCCCTGCTGCTCTGGGCAGACGATGCTGAAGATGCTCTCCTTTAAGCTGCTGCTGCTGGCCGTGGCT  
CTGGGCTTCTTTGAAAGGAGATGCTAAGTTTGGGGAAAGAAACGAAGGGAGCGGAGCAAGGAGGAGAAGGTGCCTG  
AATGGGAACCCCCGAAGCGCTGAAAGAGGAGACAGGAGGATGATGTCCCAGTGGAGCTGCTAGTGGGGGA  
GAGATGCTGTCCGGTGGCTTCTACCTCGGCTGTCTGCTGCGGAGTGACAGCCCGGGGCTAGGGCGCCTG  
GAGAATAAGATATTTTTCTGTTACCAACAACACAGAATGTGGGAAGTTACTGGAGGAAATCAATGTGCACTTTGC  
TCTCCACATTCTCAAAGCCTGTTCCACTCACCTGAGAGAGAAGTCTTGGAAAGAGACCTAGTACTTCTCTGCTC  
TGCAAAGACTATTGCAAGAATTCTTTTACACTTGCCGAGGCCATATTCCAGGTTTCTCTCAAACAACCTGCGGAT  
GAGTTTTGCTTTTACTATGCAAGAAAGATGGTGGTTGTGCTTTCCAGATTTTCCAAGAAAACAGTCAGAGGA  
CCGACATCTAACTACTTGGACCATGGAAGAATATGACAAGTGGAAAGAGATCAGCAGAAAGCAACAACAAC  
TGCTCTGTGATTACAGGAGTTGTGAGTGGGCTGCGGCAGCCCGTTGGTGCCCTGCATAGTGGGGATGGCTCGCAA  
CGTCTCTTCATTCTGAAAAAGAAGGTTATGTGAAGATACTTACCCCTGAAGGAGAAATTTCAAGGAGCCTTAT  
TTGGACATTCACAAACTTGTTCAAAGTGGAAATAAAGGGAGGAGATGAAAGAGGACTGCTAAGCCTCGCATTCCAT  
CCCAATTACAAGAAAAATGGAAAGTTGTATGTGTCTTATACCAACCAAGAACGGTGGGCTATCGGGCCTCAT  
GACCACATTCTTAGGGTTGTGAATACACAGTATCCAGAAAAATCCACCAAGTTGATTTGAGAACAGCCAGA  
GTCTTCTTGAAGTTGCAAGACTCCACAGAAAGCATCTGGGAGGACCACTGCTCTTTGGCCCTGACCGCTTTTTG  
FACATCATTCTTGGTGATGGGATGATTACACTGGATGATATGGAAGAAATGGATGGGTAAAGTGATTTACAGGC  
TCAGTGCTACGGCTGGATGTGGACACAGACATGTGCAACGTGCCTTATTCATACCAAGGAGCAACCCACACTTC  
AACAGACCAACCAGCCCCCGAAGTGTTTGCTCATGGGCTCCACGATCCAGGCAGATGTGCTGTGGATAGACAT  
CCCACTGATATAAACATCAATTTAACGATACTGTGTTCAAGTCCAATGGAAAAACAGATCATCAGCCAGAATT  
CTACAGATAATAAAGGGGAAGATATGAAGTGGCCATCACTTTAGAATTTCAAGCCATTAGTAAATGGTCTCT  
TTGGTGTGGTGGATTTGTATACCGGGGCTGCCAGTCAGAAAGATTGTATGGAAGCTACGTGTTTGGAGATCGTAAT  
GGGAATTTCTAACTCTCCAGCAAAGTCTGTGACAAAGCAGTGGCAAGAAAAACCACTCTGTCTCGGCACTAGT  
GGGTCTGTAGAGGCTACTTTTTCCGGTCACATCTTGGGATTGGAGAAGATGAAGTAGGTGAAGTTACATTTTA  
TCAAGCAGTAAAAAGTATGACCCAGACTACAATGGAAAACCTCTACAAAATTGTAGATCCCAAAAGACCTTTAATG  
CCTGAGGAATCGAGAGCCACGGTACAACCTGCACAGACACTGACTTCAGAGTGCTCCAGGCTGTGCGAAACGGC  
TACTGCACCCCCACGGGAAGTGCTGCTGCAGTCAGGCTGGGAGGGGACTCTGCAGAACTGCAAAATGTGAG  
CCAGCATGTGCTCATGGAGGTGTCTGTGTTAGACCGAACAAGTGCCTCTGTAAAAAAGGATATCTTGGTCTCAA  
TGTGAACAAGTGGACAGAAACATCCGCAGAGTGACCAGGGCAGGTATTCTTGATCAGATCATTGACATGACATCT  
TACTTGCTGGATCTAACAAGTTACATTGTATAGTCTTCTGGGACTGTTTGAATATTTCTATTCCAATGGGCATTTAT  
TTTTTATCCTGTCTATTAATAAAAAAAGACTGTTATCCTGCTACACACTCCTGTGATTTTATTCTCTTTTATTA  
TTTAAAAATAATTTCCAGAAATGTGCAGATCCTCTGTGTATGTGATGTTTGTTCACATATGCATACAC  
ATACTATAAACCCCTATATGCGTTGTTGCATAACAGATGATTTTTTAAATATATACTTCCCTTATGCAAAAGTAAT  
TTACACAGAAATTCCATTGTAAATTGATAATGGATTTTTTATGTTACTAGAAGAGATTATTTGACTTCCCAGGAA  
TTTTCTGTCTGTAATCACTAAAGTCAACTTAAATAGAGTTTTGAAACAGTACTGTGCAATCCGATGGATCTAATT  
AAAAAAAAGGCAATATTTTTATATTAAAGTACTATACTAGGAGAGATGTTTCAGAACTCCCTGATGAATTTCTA  
AGTGAGCAACTTGATATAAAATGTAATCTTCATTTTTGTGAGTGTATCCAGTTACAGAATGCTACACACTTACC  
TTTTTATTTGGCTGAGAATCTGGTTATTTTCTTAATCTCAAGATTGTTTTCAAGTGTTTTATAATTAAATCAT  
AATAGCATATTTTAAAAATCAAAA

579/615

**FIGURE 574**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139592
><subunit 1 of 1, 882 aa, 1 stop
><MW: 98428, pI: 8.89, NX(S/T): 5
MKHVTVGRNMRGCVRSCSRRRQRRPQHHLEPSAGPASACAAA VAVSVPATVPPKQLPLHG
QVNICGCPLLFFLLFQLLLLLPLPSRSRKPPTQLTLAQLQTVSSAQLYLAPRAPLRHWH
RRVFYFCKVASLYIFLSPPPPSVSGVPYSPANSSWSCALVPLLGSVPPHPPAPSPCCSG
QTMLKMLSFKLLLLAVALGFFEGDAKFGERNEGSGARRRRCLNGNPPKRLKRRDRMMSQ
LELLSGGEMLCGGFYPRLSCLLSDSPGLGRLENKIFSVTNNTECGKLEELKCALCSPH
SQSLFHSPEREVLERDLVPLLLCKDYCKEFFYTCRGHIPGFLQTTADEFCFYARKDGGL
CFPDFPRKQVRGPASNYLDQMEEDYKVEEISRKHKHNCFIQEVVSGLRQPVGALHSGDG
SQRLFILEKEGYVKILTPEGEIFKEPYLDIHKLVQSGIKGGDERGLLSLAFHPNYKKNKG
LYVSYTTNQRWAIGPHDHILRVVEYTVSRKNPHQVDLRTARVFLEVAELHRKHLGGQLL
FGPDGFLYIILGDGMITLDDMEEMDGLSDFTGSVLRDLVDVDMCNVPYSIPRSNPHFNST
NQPPPEVFAHGLHDPGRCAVDRHPTDININLTILCSDSNGKNRSSARILQIIKGKDYSEEP
SLLEFKPFSNGPLVGGFVYRGCQSERLYGSYVFGDRNGNFLTQQSPVTKQWQEKPLCLG
TSGSCRGYFSGHILGFGEDDELGEVYILSSSKSMTQTHNGKLYKIVDPKRPLMPEECRATV
QPAQTLTSECSRLCRNGYCTPTGKCCCSPGWEGDFCRTAKCEPACRHGGVCVRPNKCLCK
KGYLGPQCEQVDRNIRRVTRAGILDQIIDMTSYLLDLTSYIV
```

**Important features of the protein:****Transmembrane domains:**

Amino acids 63-80;186-201

**N-glycosylation sites:**

Amino acids 152-156;281-285;598-602;629-633;641-645

**Glycosaminoglycan attachment site:**

Amino acids 417-421

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 856-860

**N-myristoylation sites:**

Amino acids 12-18;413-419;457-463;698-695;720-726;723-729

**EGF-like domain cysteine pattern signatures:**

Amino acids 805-817;837-849

580/615

**FIGURE 575**

CGGCTCGAGAGCGGGGCAAACCTGCTTGGCACCTCTTCAATAGGTGACATTCA**ATG**ATAGATCT  
CTGGCTTCCTGCTCTGTTTGTTCTGGTTGCCCTGGAAAGCCTGCTGCTCAGCCCATGCCCCGG  
GACTTCCTCCACCCTCACCAGGACATTCTTCCATCTCTTGCTCCTGTGTGCAAGTCCCTTT  
CTCCTGGATTCCATGTCTTGAATGTTTCTTAATTTACTTCCTCATTTTGGCAGAGGATGTCTT  
CCAGTTGTTTTCTGGGAATGCTAATATGCAAGTGAACCAG**TGA**CCTGCAGTTCTGCCACACA  
GGGTTAATAACCAATCAGATTCTCTCTTTTCAAGATGGTTAACATAACAGACACCAAGAAAGG  
GAAGAGGAGCCGACAGCAGAGGGGGAAGCTGAAAAGACGCACAAAGAATGGCCATAAAAGATA  
TGAGCAACCCAGCTTTCCAGACAGTCACTTTTCCCAGTGGTCATACCTGGTCTGGAAGATTC  
CCCATCATCTCGAATAAAGCTGTTGTTGCTTTTAACTCCATGGAGAGACCGAATGGAGTGAGC  
CCAGCAGGGCATGCTGGGCAAGAGAGGTCCCCGAGTCCCAAATAAGAATTTCAACTAGTATA  
AAACGAGGCAGCGAACCCACACGTGGAAGTCTGATACCGCTTGCAAGAGGGAATTGAATAGAT  
GTCTCCCTATTGGTAAGGATGTGGTTTTTATTGACTTGAAATAACAAAGCCCGCAAGCAACAAC  
TGATCATCCGCGGGATGCTGCCACAAGGAATAATTGAGCACTCATTCAGACACAGGGGAAACC  
ACTGCCTCTTTCAGTCTTTCTCCCAGATTCCAACAGTCAGTGTTACAGCATTTTACCTTGTTT  
ACCTCCCTGAGAAGACGTTGCAGCTCACTACCCAGTGGGCACTGGGAGCCTCTGCTCAGGT  
GGGAGACAGATGCCCCCACATGCACATCTGGTGTATTATGAAGCAGATACTGGGGCTTCATAA  
ACACAGAAGGGGCAGGGAAGTAGCCCCAGGGCATAGTGTGGGGCCTCCTGACTAAAAGTAGCTT  
GCAAACCCCTGCCTATAACAGCCACTTCCTGGCAGTCATTGTGCCACTTAGGAGCCCTCCTCA  
CCGCTCTTCCCTTTAGTTCCTTCACTCTGTATAGACCCTGCCAGAGCAGCTCAGGGTGGAGCAG  
CTGCAGCCATGGGACCTGCTCCAGGCAAGGCCCTATGCTACACAGTCCCTGGGGGTGAGGATT  
CCAGAGGAGTCAATTCCCTACCCGCCCTTCCAGGGGATGGCCGAGATGAGCGTTCCACAGGG  
AAAGTGAAAGTTGCAAGGCTGTTGCGGAAACACCAGGGGTTCATTTAGGTCTTGCTGCTCAT  
CTCACAGAAAGCCAGTCACTGAGACAAGTATTGTGAGGGAAGAAGGCTTTATCCAGGTGCTAC  
AGACAGGGTGAACAAGAGATCAGTCTCAAATCCATCATCAACTGACTAAAATTAGGGGTT  
TATATAGCAGGGAAGAAATGTAACATACATGTGGGAAAACAGGGATTAACGAGGGGCAAGGAAG  
AGGAGTTGGCCAACAGGCAGCAGGTGGTCACCTGGGGAATCATGATGGGTGAGGGGTCTGGCT  
TCTCACTTTCCAGATGTGGGGATCTGGTAAATTTTCACTTTCTTGATACCGTCTGGGAGGATTG  
CTGGCTGGTTTCTGAGAAAGGAATTCAGATGACATAAATGTAAATTTCTCCTTGGGGTTTCAA  
GACTGAGAGGGTCAATTTCTAGGTTTATTCAAGAAAAACCATAAACATCAGTTCTATGGGACA  
ATTGGGCCCATTTCAGGCTCTGAGGATAAGGGTTAATGAGGGGACAGAGTCGCTGGAGAAG  
TTCAGTGGGGCCTACAAGAACTAGAGAGGCTTCCTGGCAAAGCTCTATGCTGTCTATCCTCT  
CTTCTCTCCTTGACAGGAAGATTCCAGTATAATAGACCCGAGGTGAAAAGGCTTTTGTTCATA  
AGTAGAAAACCTGAAGGGGGGTGGGAGGCACATGGATTTGAACCAGAGACCGCTTGGCCTGCAG  
GCAAATGCTGTACCTTCAGTTGCACCCCTCACTTGTTACAGCTGTTTCTGATAAGCACTTGTTG  
CAGCCCCATCAGCACCTCGATTTCTTCTTGGTGAGTCCATGGGAACAGCCCCACTGCAAACAA  
CCCATTCTGCTCTCCTCTTTCTTAAACCTCAACCTCCTCCTACCTGGCAGTCCACAGGCC  
TACAGCTTCTCCTCAGTGGGAAAGACATCAGCTTGGAACCACTTGGAAGCCAACGTTATC  
CTAGAAAAGCTTTTTTAAATGACCCAGCAGGACAAGTCTCCGGATGGCCTTGGCCAACCCGGTG  
CTTCCCTCTTTCTTGGTTGTAGTTCTCAGAATAACTAGAGAATGTACTGGGAGTGTTGTCTCT  
GAGATAAGGAGGAACTGTCCTAAACCTGGACTCTGTTCCCATCACACCTAGAACAGGATGTCC  
TGCAACGCTTTAGCCCAATGATCCAAGTTGCCCTTGGGGTATAAACTTGACAGCAGAGGGCG  
TTCAGGGTCCCTCAGCTGCAGTGTGAAGTGGGACACACAGGTGAGACTCCATCTGCCCTGGGC  
AGGTTCTGAGCCTTGGGGGACCAGTTTACCCTACATCCCAGGCTTCTGTTGTCCCTTGCTG  
CCTGTAAGGAATAAAGTTGCTTTGCTTA

581/615

**FIGURE 576**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA139608
><subunit 1 of 1, 80 aa, 1 stop
><MW: 8927, pI: 3.77, NX(S/T): 0
MIDLWLPALFVLVALESLLLSPCPGTSSTLTRTFFPSLVSCVQVPFSWIPCLECFLIYFL
ILAEDVLQLFSGNANMQVNQ
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-29

**Transmembrane domain:**

Amino acids 47-62

**N-myristoylation sites:**

Amino acids 25-31



582/615

**FIGURE 577**

ATCGGTTAGCGCCTTGCCATGATTAATCCAGAGCTGCGGGATGGCAGAGCTGATGGCTTCATA  
CATCGGATAGTTCCCAAGTTGATACAAACTGGAAGATTGGCCTTATGTGCTTCCTGAGTATT  
ATTATTACTACAGTTTGCATTATTATGATAGCCACATGGTCCAAGCATGCTAAACCTGTGGCA  
TGTTCAGGGGACTGGCTTGGAGTGAGAGATAAGTGTTTCTATTTTCTGATGATACCAGAAAT  
TGGACAGCCAGTAAAATATTTTGTAGTTTGCAGAAAGCAGAACTTGCTCAGATTGATACACAA  
GAAGACATGGAATTTTGAAGAGGTACGCAGGAAGTATGCACTGGATTGGACTAAGCAGG  
AAACAAGGAGATTCTTGGAATGGACAAATGGCACCACATTCAATGGTTGGCCATCAAACCTCC  
AAATGGTCTTGCAACTGGAGCCTCCGACAATGGCTTCTTCTGCTGGGACCCCTTAGATTAGGCC  
TCTGAGGGAGCTCTGACTGCCGTTTCCCCAAAACAATGTCCCCTGTCAGCAGGAAGCAGTTAA  
ATCAGTCTTCATCCTTATCCTTAATATAACGGCAGTTAGATGTACTTCTTTAGAGGGAGTAAA  
TTTATCAATTCAGAGCAATTCATCCTCCTCTTTCCATCTTTGATTACAGTTAATAGGCTATA  
AATTTTGATAATGTAGAATAAACTACAGAAAACCTTCTTG

583/615

**FIGURE 578**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA143292
><subunit 1 of 1, 160 aa, 1 stop
><MW: 18576, pI: 9.29, NX(S/T): 3
MINPELRDGRADGFIHRIVPKLIQNWKIGLMCFLSIIITTVCIIIMIATWSKHAKPVACSG
DWLGVRDKCFYFSDTRNWTASKIFCSLQKAELAQIDTQEDMEFLKRYAGTDMHWIGLSR
KQGDSWKWTNGTTFNGWPSNSKWSCNWSLRQWLLLLGPLR
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-42

**N-glycosylation sites:**

Amino acids 78-82;130-134;146-150

**N-myristoylation site:**

Amino acids 131-137

584/615

**FIGURE 579**

TGAAGGCCTGTGAGTGAGGAATGCCTCTCACCAGCTGTGCCTGAGCTGCAGCACTCCAGCCAC  
TGCTGTCTCCTTAGCTGCTCACAT**ATGG**ATACTTTCACAGTTCAGGATTCCACTGCAATGAGC  
TGGTGGAGGAATAATTTCTGGATCATCTTAGCTGTGGCCATCATTGTTGTCTCTGTGGGCCTG  
GGCCTCATCCTGTACTGTGTCTGTAAGTGGCAGCTTAGACGAGGCAAGAAATGGGAAATTGCC  
AAGCCCCTGAAACACAAGCAAGTAGATGAAGAAAAGATGTATGAGAATGTTCTTAATGAGTCG  
CCAGTTCAATTACCGCCTCTGCCACCGAGGAATTGGCCTTCTCTAGAAGACTCTTCCCCACAG  
GAAGCCCCAAGTCAGCCGCCCCGCTACATACTCACTGGTAAATAAAGTTAAAAATAAGAAGACT  
GTTTCCATCCCAAGCTACATTGAGCCTGAAGATGACTATGACGATGTTGAAATCCCTGCAAAT  
ACTGAAAAAGCATCATTTT**TGAA**ACAGCCATTTCTTCTTTTTGGCAAAACTGAAGAGGGTTCAC  
ACAACCTATTTTAAAAACAATCAAGAATGGTTGAACTTCAGTAGGTCTCTGGGCCCTGAAAGCC  
AGTGGTGATTTTATGAAGCTCTATAAGATAAAGCACTTCCCAAACCTTAGATGAAGACACCCC  
TGCGATCGGATGACTGCAGCCAGAGGAGACACATGGGTGCTCGGCTCTGAGGACTTAGAGGGG  
TCAGCCTTGTGCTGTTGAGGAACTTTCCATGGGAAGGACCACGGGGCTCCATGGCTCCCACC  
TGTGGGAACTACTCATTTCTTGGCATTCTTTCCCCCTTCATTCCCTTTGGTTTGCATGGTTC  
TGAGTGATATTAAATCTCAGCATTTGGTTGTGCAAAAAAAAAA

585/615

**FIGURE 580**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA144844
><subunit 1 of 1, 145 aa, 1 stop
><MW: 16618, pI: 5.26, NX(S/T): 1
MDTFTVQDSTAMSWWRNFWIILAVAIIVSVGLGLILYCVCKWQLRRGKKWEIAKPLKH
KQVDEEKMYENVLNESPVQLPPLPPRNWPSLEDSSPQEAPSQPPATYSLVNKVKNKKTVS
IPSYIEPEDDYDDVEIPANTEKASF
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-35

**Tyrosine kinase phosphorylation site:**

Amino acids 61-70

**Amidation site:**

Amino acids 48-52

586/615

**FIGURE 581**

GGCCGCCTCCGCGGGGCTGTGGGAAGCTTGGGCTGTCCCAGGACCGTCAGTCTCCTCCTCTGA  
CCCTCCCTTTCCCCTTGTGTGTAGGGCCGCGTCCCACCCCCACCTCGCCGGAGTCCGGGGCG  
GCCCCGGTGTCCCCTCCGAGCCTGCTGCACTCCACGTCCCCCTACCAGGGCTCCAGCCCCAG  
GGAAATCTCCGACCAGGCCCGCCAGGAGCCAGATCCAGGCTCCTGGAAGAACCATGTCCGGC  
AGCTACTGGTCATGCCAGGCACACACTGCTGCCCCAAGAGGAGCTGCTGTTTGAATTATCTGTG  
AATGTTGGGAAGAGGAATGCCAGAGCTGCCGGCTGAAAATTACCCAACCAAGAGAAATCTGCAGG  
**ATG**GACTTTCTGGTCCTCTTCTTGTCTACCTGGCTTCGGTGCTGATGGGTCTTGTCTTATC  
TGCGTCTGCTCGAAAACCCATAGCTTGAAAGGCCTGGCCAGGGGAGGAGCACAGATATTTTCC  
TGTATAATTCCAGAATGTCTTCAGAGAGCCGTGCATGGATTGCTTCATTACCTTTTCCATACG  
AGAAACCACACCTTCATTGTCTGCACCTGGTCTTGCAAGGGATGGTTTATACTGAGTACACC  
TGGGAAGTATTTGGCTACTGTCAGGAGCTGGAGTTGTCTTGCAATTACCTTCTTCTGCCCTAT  
CTGCTGCTAGGTGTAAACCTGTTTTTTTTTACCCTGACTTGTGGAACCAATCCTGGCATTATA  
ACAAAAGCAAATGAATTATTATTTCTTCATGTTTATGAATTTGATGAAGTGATGTTTCCAAAG  
AACGTGAGGTGCTCTACTTGTGATTTAAGGAAACCAGCTCGATCCAAGCACTGCAGTGTGTGT  
AACTGGTGTGTGCACCGTTTTCGACCATCACTGTGTTTGGGTGAACAACATGCATCGGGCCTGG  
AACATCAGGTACTTCCTCATCTACGTCTTGACCTTGACGGCCTCGGCTGCCACCGTCGCCATT  
GTGAGCACCCTTTTCTGGTCCACTTGGTGGTGATGTCAGATTATACCAGGAGACTTACATC  
GATGACCTTGGACACCTCCATGTTATGGACACGGTCTTTCTTATTCAGTACCTGTTTCTGACT  
TTTCCACGGATTGTCTTCATGCTGGGCTTTGTGCTGGTCTGAGCTTCCTCCTGGGTGGCTAC  
CTGTTGTTTGTCTGTATCTGGCGGCCACCAACCAGACTACTAACGAGTGGTACAGAGGTGAC  
TGGGCCTGGTGCCAGCGTTGTCCCCTTGTGGCCTGGCCTCCGTCAGCAGAGCCCCAAGTCCAC  
CGGAACATTCACTCCCATGGGCTTCGGAGCAACCTTCAAGAGATCTTTCTACCTGCCTTTCCA  
TGTCATGAGAGGAAGAAACAAGAA**TGA**CAAGTGTATGACTGCCTTTG

587/615

**FIGURE 582**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA144857
><subunit 1 of 1, 344 aa, 1 stop
><MW: 39787, pI: 7.44, NX(S/T): 2
MDFLVLFLEFYLASVLMGLVLICVCSKTHSLKGLARGGAQIFSCIIPECLQRAVHGLLHLYL
FHTRNHTFIVLHLVLQGMVYTEYTWEVFGYCQELELSLHYLLLPYLLLGVNLEFFFTLTCTG
TNPGIITKANELLFLHVYEFDEVMFPPKNVRCSTCDLRKPARSKHCSVCNWCVHRFDHHCV
WVNNCIGAWNIRYFLIYVLTLTASAATVAIVSTTFLVHLVVMSDLYQETYIDDLGHLHVM
DTVFLIQYLFLTFFPRIVFMLGFVVVLSFLLGGYLLFVLYLAATNQTTNEWYRGDWAWCQR
CPLVAWPPSAEPQVHRNIHSHGLRSNLQEIFLPAFPCHERKKQE
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-29

**Transmembrane domains:**

Amino acids 100-116;201-217;256-275

**N-glycosylation sites:**

Amino acids 65-69;284-290

**N-myristoylation sites:**

Amino acids 32-38;77-83;120-126;322-328

**Cell attachment sequence:**

Amino acids 292-298

**DHHC zinc finger domain:**

Amino acids 140-204

588/615

**FIGURE 583**

CCGCGGAAC TGGCAGGCGTTTCAGAGCGTCAGAGGCTGCGGATGAGCAGACTTGGAGGACTCCAGGCCAGAGACT  
AGGCTGGGCGAAGAGTTCGAGCGTGAAGGGGGCTCCGGGCCAGGGTGACAGGAGGCGTGCTTGAGAGGAAGAAGTT  
GACGGGAAGGCCAGTGGCAGGGCAAATCTCGTGAACCTTGGGGGACGAATGCTCAGGATGCGGGTCCCCGCCCTC  
CTCGTCTCTCTCTCTGCTTCAGAGGGAGAGCAGGCCCCGTCGCCCCATTTCTGCAACAGCCAGAGGACCTGGTG  
GTGCTGCTGGGGGAGGAAGCCCCGCTGCCGTGTCTCTGGGCGCCTACTGGGGGCTAGTTCACTGGACTAAGAGTGGG  
CTGGCCCTAGGGGGCCAAAGGGACCTACCAGGGTGGTCCCGGTACTGGATATCAGGGAATGCAGCCAATGGCCAG  
CATGACCTCCACATTAGGCCCCGTGGAGCTAGAGGATGAAGCATCATATGAATGTAGGCTACACAAGCAGGCGCTC  
CGCTCCAGACCAGCCCAACTGCACGTGCTGGTCCCCCAGAAAGCCCCCAGGTGCTGGGCGGCCCCCTCTGTGTCT  
CTGGTTGCTGGAGTTCTTGCAGACCTGACATGTGGAGCCGTGGGGATGCCCGCCCTACCCCTGAATTGCTGTGG  
TTCCGAGATGGGGTCTGTTGGATGGAGCCACCTTTTCATCAGACCTGCTGAAGGAAGGGACCCCTGGGTCACTG  
GAGAGCACCTTAACCCCTGACCCCTTTTCAGCCATGATGATGGAGCCACCTTTGTCTGCCGGGCCCCGAGCCAGGCC  
CTGCCCACAGGAAGAGACACAGCTATCACACTGAGCCTGCAGTACCCCCAGAGGTGACTCTGTCTGCTTCGCCA  
CACACTGTGCAGGAGGAGAGAAGGTCATTTTCTGTGCCAGGCCACAGCCAGCCTCCTGTACAGGCTACAGG  
TGGGCAAAGGGGGCTCTCCGGTGCTCGGGGCCCGGGGCCAAGGTTAGAGGTCGTGGCAGACGCTCGTTCTCTG  
ACTGAGCCCGTGTCTGCGAGGTGAGCAACGCCGTGGGTAGCGCCAACCGCAGTACTGCGCTGGATGTGCTGTTT  
GGGCGGATTCTGCAGGCAAAGCCGGAGCCCGTGTCCGTGGACGTGGGGGAAGACGCTTCTTCAGTGCAGCTGG  
CGCGGGAACCCGCTTCCACGGTAACCTGGACCCGCCGCGGTGGCGCGCAGGTGCTGGGCTCTGGAGCCACACTG  
CGTCTTCCGTGGTGGGGCCCCGAGGACGAGGCGACTATGTGTGAGAGCTGAGGCTGGGCTATCGGGCTGCGG  
GGCGGCGCGCGGAGGCTCGGCTGACTGTGAACGCTCCCCAGTAGTGACCGCCCTGCAGTCTGCGCTGCGCTTC  
CTGAGGGGCCCTGCTCGCCTCCAGTGTCTGGTTTTTCGCCCTCTCCCGCCAGATGCCGTGGTCTGGTCTTGGGAT  
GAGGGCTTCTGGAGGCGGGGTGCGAGGGCCGGTTCTGGTGGAGACATTCCCTGCCCCAGAGAGCCGCGGGGGA  
CTGGGTCCGGGCTGATCTCTGTGCTACACATTTCCGGGACCCAGGAGTCTGACTTTAGCAGGAGCTTTAACTGC  
AGTGGCCGGAACCGCTGGGCGAGGGAGGTGCCAGGCCAGCCTGGGCCGTAGAGACTTGCTGCCCAGTGTGCGG  
ATAGTGGCCGGAGTGGCCGCTGCCACCACAACCTCTCCTTATGGTCACTACTGGGGTGGCCCTCTGCTGCTGGCGC  
CACAGCAAGGCCTCAGCCTCTTCTCCGAGCAAAGAACCTGATGCGAATCCCTGGCAGCAGCGACGGCTCCAGT  
TCACGAGGTCTGAAGAAGAGGAGACAGGCAGCCGCGAGGACCGGGGCCCATTTGTGCACACTGACCACAGTGAT  
CTGGTTCTGGAGGAGGAAGGGACTCTGGAGACCAAGGACCCAACCAACGGTTACTACAAGGTCCGAGGAGTCAGT  
GTGAGCCTGAGCCTTGGCGAAGCCCTGGAGGAGGTCTCTTCTGCCACCACCCCTCCCCCTTGGGCCCCCAGG  
ACCCCTACCTTCTATGACTTCAACCCACACCTGGGCATGGTCCCCCCTGCAGACTTTACAGAGCCAGGGCAGGC  
TATCTCACCACACCCACCCCTCGAGCTTTCACCAGCTACATCAAACCCACATCCTTTGGGCCCCAGATCTGGCC  
CCCGGGAATCCCCCTTCCCATATGCTGCCTTCCCCACACCTAGCCACCCGCGTCTCCAGACTCAGTGTGACAT  
CTTTCCAATGGAAGAGTCTCTGGGATCTCCAACCTGCCATAATGGATTGTTCTGATTTCTGAGGAGCCAGGACAAG  
TTGGCGACCTTACTCTCCAAAACCTGAACACAAGGGGAGGGAAAGATCATTACATTTGTGAGGAGCATTTGTATA  
CAGTCAGCTCAGCCAAAGGAGATGCCCCAAGTGGGAGCAACATGGCCACCCAATATGCCACCTATTCCCCGGTG  
TAAAAGAGATTCAAGATGGCAGGTAGGCCCTTTGAGGAGAGATGGGGACAGGGCAGTGGGTGTTGGGAGTTTGGG  
GCGGGGATGGAAGTTGTTTCTAGCCACTGAAAGAAGATATTTCAAGATGACCATCTGCATTGAGAGGAAAGGTAG  
CATAGGATAGATGAAGATGAAGAGCATACCAGGCCCCACCCCTGGCTCTCCCTGAGGGGAACCTTGCTCGGCCAAT  
GGAATGCAGCCAAGATGGCCATATACTCCCTAGGAACCCAAAATGGCCACCATCTTGATTTTACTTTCTTAAA  
GACTCAGAAAGACTTGGACCAAGGAGTGGGGATACAGTGAGAATTACCACTGTTGGGGCAAATATTGGGATAA  
AAATATTTATGTTTAAATAATAAAAAAAGTCAAAGAGAAAAAAA

589/615

**FIGURE 584**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA145841
><subunit 1 of 1, 708 aa, 1 stop
><MW: 75093, pI: 6.65, NX(S/T): 3
MLRMRVPALLVLLFCFRGRAGPSPHFLQQPEDLVVLLGEEARLPCALGAYWGLVQWTKSG
IALGGQRDLPGWSRYWISGNAANGQHDHLIRPVELEDEASYECQATQAGLRSRPAQLHVL
VPPEAPQVLGGPSVSLVAGVPANLTCRSRGDARPTPELLWFRDGVLLDGATFHQTLLKEG
TPGSVESTLTLPFSHDDGATFVCRARSQALPTGRDTAITLSLQYPPEVTLSASPHTVQE
GEKVIFLCQATAQPPVTGYRWAKGGSPVLGARGPRLEVVDASFLTEPVSCVSNVAVGSA
NRSTALDVLFGPILQAKPEPVSVVDGEDASFSCAWRGNPLPRVTWTRRGGAQVLGSGATL
RLPSVGPEDAGDYVCRAEAGLSGLRGGAEARLTVNAPPVVVTAHSAFAFLRGPARLQCL
VFASAPDAVWVSWDEGFLEAGSQGRFLVETFPAPESRGGLGPGGLISVLHISGTQESDFS
RSFNCSARNRLGEGGAQASLGRRDLLPTVRIVAGVAAATTTLLMVITGVALCCWRHASKAS
ASFSEQKNLMRIPGSSDGSSSRGPEEEEETGSREDRGPIVHTDHSDLVLEEEGTLETKDPT
NGYYKVRGVSVSLSLGEAPGGGLFLPPPSPLGPPGTPTFYDFNPHLGMVPPCRLYRARAG
YLTPHPRAFTSYIKPTSFSGPPDLAPGTPPPFYAAFPPTSPHRLQTHV
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**Transmembrane domain:**

Amino acids 511-531

**N-glycosylation sites:**

Amino acids 143-147;301-305;484-488

**N-myristoylation sites:**

Amino acids 48-54;60-66;79-85;139-145;180-186;183-189;355-361;383-389;  
387-393;460-466;473-479;494-500;495-501;514-520;528-534;  
554-560;592-598;608-614

**Amidation site:**

Amino acids 500-504

**Cell attachment sequence:**

Amino acids 149-152

**Multicopper oxidases signature 1:**

Amino acids 445-466

**Immunoglobulin domain:**

Amino acids 326-377



590/615

**FIGURE 585**

GCCCGCCTGAGGAAGCCGTGTGCCTGGGATGCCAAGAGCCAGAGAATGGATCTTCTCCGAGTG  
GGGACATTGCTGACAATCCCGGCTTCCCGAGGCGGCTAAGAACAGGCAGTTTGTGTGCGGCTGG  
CTGCAGATAACCCAGAGGCACAAAGAGACCGAAGCCACCCGGAGGGACCCACGGACGGACAGAT  
GGTAGGCGCGAACCCGAGAGGACCGGCGGAGGCTGAGCACCGAGAGCCGCCAAGGAAGAGAAA  
CTAACACAGCCAAGTTACCCCGCGGCTTTCCTTCGCTGCGCTAAGGAATGAAACCCCTTCCA  
GCTCGATCTGCTCTTCGTCTGCTTCTTCTCTTCAGTCAAGAGCTGGGCTCCAGAAGAGAGG  
ATGCTGTCTGGTGTGCTGGGCTACATGGCCAAGGACAAGTTTCGGAGAATGAATGAAGGCCAAGT  
CTATTTCCTTCAGCCAGCAGCCCCAGGACCAGGTGGTGGTGTGCGGACAGCCAGTGACGCTACT  
TTGCGCCATCCCCGAATACGATGGCTTCGTTCTGTGGATCAAGGACGGCTTGGCTCTGGGTGT  
GGGCAGGGACCTCTCAAGTTACCCACAGTACCTGGTGGTAGGGAACCACTGTGAGGGGAGCA  
CCACCTGAAGATCCTGAGGGCAGAGCTGCAAGACGATGCGGTGTACGAGTGCCAGGCCATCCA  
GGCCGCCATCCGCTCCCGCCCCGCACGCCTCACAGTCTGGTGCCGCCTGATGACCCCGTCAT  
CCTGGGGGGCCCTGTGATCAGCCTGCGTGCGGGGGACCCCTCTCAACCTCACCTGCCACGCAGA  
CAATGCCAAGCCTGCAGCCTCCATCATCTGTTGCGAAAGGGAGAGGTTCATCAATGGGGCCAC  
CTACTCCAAGACCCTGCTTCGGGACGGCAAGCGGGAGAGCATCGTCAGCACCCCTCTTCATCTC  
CCCTGGTGACGTGGAGAATGGCCAGAGCATCGTGTGTGCTGCCACCAACAAAGCCATCCCCGG  
AGGAAAGGAGACGTCGGTCACCATTCAGATCCAGCACCCCTCCACTGGTCAACCTCTCGGTGGA  
GCCACAGCCAGTGCTGGAGGACAACGTCGTCACCTTCCACTGCTCTGCAAAGGCCAACCCAGC  
TGTCACCCAGTACAGGTGGGCCAAGCGGGGCCAGATCATCAAGGAGGCATCTGGAGAGGTGTA  
CAGGACCACAGTGGACTACACGTAATCTCAGAGCCCGTCTCCTGTGAGGTGACCAACGCCCTG  
GGCAGCACCAACCTCAGCCGCACGGTTGACGTCTACTTTGGGCCCCGGATGACCACAGAACCC  
CAATCCTTGCTCGTGGATCTGGGCTCTGATGCCATCTTCAGCTGCGCCTGGACCGGCCAACCCA  
TCCCTGACCATCGTCTGGATGAAGCGGGGCTCCGGAGTGGTCTGAGCAATGAGAAGACCCTG  
ACCCTCAAATCCGTGCGCCAGGAGGACGCGGGCAAGTACGTGTGCCGGGCTGTGGTGCCCCGT  
GTGGGAGCCGGGGAGAGAGAGGTGACCCTGACCGTCAATGGACCCCCATCATCTCCAGCACC  
CAGACCCAGCACGCCCTCCACGGCGAGAAGGGCCAGATCAAGTGCTTCATCCGGAGCACGCCG  
CCGCCGGACCGCATCGCCTGGTCTGGAAGGAGAACGTTCTGGAGTCGGGCACATCGGGGCGC  
TATACGGTGGAGACCATCAGCACCGAGGAGGGCGTCATCTCCACCCTGACCATCAGCAACATC  
GTGCGGGCCGACTTCCAGACCATCTACAACCTGCACGGCCTGGAACAGCTTCGGCTCCGACACT  
GAGATCATCCGGCTCAAGGAGCAAGGTTTCGAAATGAAGTCGGGAGCCGGGCTGGAAGCAGAG  
TCTGTGCCGATGGCCGTATCATTTGGGGTGGCCGTAGGAGCTGGTGTGGCCTTCTCGTCCTT  
ATGGCAACCATCGTGGCGTTCTGCTGTGCCCGTTCCAGAGAAGTACGGGAGGGAGATCCGGG  
ATCTCAGGGAGGGGGACAGAGAAAAGGCCAGGCTTAGGCTGCCCCGGAGAGCAAGTAAGCAG  
GAGTGCAATGAACAGGGGTCC**TAA**CAGTGCTGTGAGCTCCTGGGGCAGGGAGTGGGTCTGATG  
CATCGGTGTATGTGAGCCTGGGCAACATGGCGCCTGGCAGAGTGGGCGCTAGGCTGAGGTGTA  
CCTGGACTAGACTGAACTTCATCTGCAGGGCAGCCAGCATTTTGGATTGAACACATAGCTCTT  
TCAGTCAGGAACTGTACAGAAAGATAGGGGGAAAAGCGGTTTGTGGTTTGATCCTTGCTCTAC  
AAGAGCTGTTAGTCTAGAGAGACCCCATCTCTACAACAAAATAAAAATAAAGAGCTGCTAGTC  
TCACCAGAAAAGCAGGTCACTCACACAGCTGTGGGGGAGTGGGTGGGGAAGCAATAAAGGAAT  
TGCTTTGAGAAAACCTTA

591/615

**FIGURE 586**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA148004
><subunit 1 of 1, 600 aa, 1 stop
><MW: 65308, pI: 8.35, NX(S/T): 5
MKPFQLDLLFVCFFLFSQELGLQKRGCCLVGYMAKDKFRRMNEGQVYSFSQQPQDQVVV
SGQPVTLLCAIPEYDGFVLWIKDGLALGVGRDLSSYPQYLVVGNHLSGEHHLKILRAELQ
DDAVYECQAIQAAIRSRPARLTVLVPPDDPVILGGPVISLRAGDPLNLTCHADNAKPAAS
IIWLRKGEVINGATYSKTLLRDGKRESIVSTLFISPGDVENGQSIVCRATNKAIPGGKET
SVTIDIQHPPLVNLSVEPQPVLEDNVVTFHCSAKANPAVTQYRWAKRGQIIKEASGEVYR
TTVDYTYFSEPVSCVTNALGSTNLSRTVDVYFGPRMTTEPQSLLVDLGSDAIFSCAWTG
NPSLTIVWMKRGSGVVLNSNEKTLTLKSVRQEDAGKYVCRAVVPRVGAGEREVTTLTVNGPP
IISSTQTQHALHGEKGQIKCFIRSTPPPDRIAWSWKENVLESSTSGRYTVETISTEEGVI
STLTISNIVRADFQTIYNCTAWNSFGSDTEIIRLKEQGSEMKSGAGLEAESVPMAVIIGV
AVGAGVAFLVLMATIVAFCCARSQRSTGGRSGISGRGTEKKARLRLPRRASKQECNEQGS
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-17

**Transmembrane domain:**

Amino acids 534-555

**N-glycosylation sites:**

Amino acids 167-171;253-257;324-328;498-502

**Glycosaminoglycan attachment sites:**

Amino acids 523-527;574-578

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 204-208;370-374;588-592

**Tyrosine kinase phosphorylation sites:**

Amino acids 40-49;300-308;389-397

**N-myristoylation sites:**Amino acids 45-51;62-68;84-90;103-109;192-198;236-242;  
374-380;436-442;478-484;539-545;543-549;  
568-574**Amidation site:**

Amino acids 202-206

**Leucine zipper pattern:**

Amino acids 8-30

592/615

**FIGURE 587**

CAAAAAAGGAGCATGTCTTCATCCATGAGAGGCCTCGAAAAATCAATGGTATTTGCATTTCTC  
CCAAGAAGGTTGCTTGCCAAAACCTTTTCGGCCATTTTCTGCTTTCAGAGTGAGACAAAGTTCA  
AAATGACAGTCTGTGAGCTCATTGAAGGCACTAGATACCCTGCCTGCAGGTACCACTATTCCC  
CCACAGAGGGGTTTGTCTTGTCACTTGTGATGACTTGAGGCCAGATAGTTTCCTTGGCTATG  
TTAAATAACTCAAGATCAGCTACCGAGTCTGAGATCTCTTCTCTCATGGCATTGGAGCTGGCT  
GTGCCTGAGGCAGACCTGGACCGTGGACATGGGGCAATGCCTTGAGCGGAAGGGGAAGCCACT  
GAATTTTGGGTGTCACCAGGTAAACAGAGCCCTCAGCATCTGAATAGAACTGAACAGGAACA  
GAAGAGATTACACTACATCTGAG**ATG**GAGACCTTTCCTCTGCTGCTGCTCAGCCTGGGCCTGG  
TTCTTGCAGAAGCATCAGAAAGCACAATGAAGATAATTAAAGAAGAATTTACAGACGAAGAGA  
TGCAATATGACATGGCAAAAAGTGGCCAAGAAAAACAGACCATTGAGATATTAATGAACCCGA  
TCCTGTTAGTTAAAAATACCAGCCTCAGCATGTCCAAGGATGATATGTCTTCCACATTACTGA  
CATTCAGAAGTTTACATTATAATGACCCCAAGGGAAACAGTTCGGGTAATGACAAAGAGTGTT  
GCAATGACATGACAGTCTGGAGAAAAGTTTCAGAAGCAAACGGATCGTGCAAGTGGAGCAATA  
ACTTCATCCGCAGCTCCACAGAAGTGATGCGCAGGGTCCACAGGGCCCCCAGCTGCAAGTTTGTA  
CAGAATCCTGGCATAAGCTGCTGTGAGAGCCTAGAACTGGAAAATACAGTGTGCCAGTTCCT  
ACAGGCAAACAATTCCCCAGGTGCCAATACCATAGTGTTACCTCATTAGAGAAGATATTGACA  
GTGCTGACAGGTCACTCTCTGATGAGCTGGTTAGTTTGTGGCTCTAAGTTG**TAA**ATCCACAG  
AGCTTTAGGACTAGGGTCTTACTAAAGAAGGACCTCTTCTTGTTCATTCTTGTTTAAACCTTT  
CCTTAATATCTACTCTTTAGCACTATAGTGAACCTCTGATTATTTATTCTAACTGGAGGAGTG  
AAAAATCCAAAATTGTGGATAATTCAATTAAAAGTTATGACTGATACCG

593/615

**FIGURE 588**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA149893
><subunit 1 of 1, 199 aa, 1 stop
><MW: 22427, pI: 6.46, NX(S/T): 3
METFPLLLLSLGLVLAEESESTMKIIKEEFTDEEMQYDMAKSGQEKQTIEILMNPILLVK
NTSLMSKDDMSSTLLTFRSLHYNDPKGNSGNDKECCNDMTVWRKVSEANGSCKWSNNF
IRSSTEVMRRVHRAPSCKFVQNPGISCCESLELENTVCQFTTGKQFPRCQYHSVTSLEKI
LTVLTGHSLMSWLVCGSKL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-16

**N-glycosylation sites:**

Amino acids 61-65;89-93;111-115

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 105-109

**N-myristoylation sites:**

Amino acids 12-18;88-94;144-150

**Microbodies C-terminal targeting signal:**

Amino acids 197-201

594/615

**FIGURE 589**

CAGTCCTGCCGGGACGGTGAGCGCATTACAGCACCCCTGGACAGCACCGCGGTTGCGCTGCCTCC  
AGGGCGGCCCCGGGCTGCTCCTGCTCCGCAGAGCTACGCCCTCCCCCGGGTGCCCCGGACCC  
TGCACTTGCCGCCGCTTTCCCTCGCGCTGCTCTGGACCTTGCTAGCCGGCTCTGCACCTCCAG  
AAGCCGTGGGCGCGCCGCTCAGCTGCTCCATCGCCTCACTTTCCCAGGCTCGCGCCCGAAGCA  
GAGCCATGAGAACCCAGGGTGCCTGGCGAGCCGCTAGCGCC**ATGGG**CCCCGGCGAGGCGCTG  
CTGGCGGGTCTCCTGGTGATGGTACTGGCCGTGGCGCTGCTATCCAACGCACTGGTGCTGCTT  
TGTTGCGCCTACAGCGCTGAGCTCCGCACTCGAGCCTCAGGCGTCCTCCTGGTGAATCTGTCT  
CTGGGCCACCTGCTGCTGGCGGCGCTGGACATGCCCTTCACGCTGCTCGGTGTGATGCGCGGG  
CGGACACCGTCGGCGCCCCGGCGCATGCCAAGTCATTGGCTTCCTGGACACCTTCCTGGCGTCC  
AACGCGGCGCTGAGCGTGGCGGCGCTGAGCGCAGACCAGTGGCTGGCAGTGGGCTTCCCACTG  
CGCTACGCCGGACGCCTGCGACCGCGCTATGCCGGCCTGCTGCTGGGCTGTGCCTGGGGACAG  
TCGCTGGCCTTCTCAGGCGCTGCACTTGGCTGCTCGTGGCTTGGCTACAGCAGCGCCTTCGCG  
TCCTGTTTCGCTGCGCCTGCCGCCGAGCCTGAGCGTCCGCGCTTCGCAGCCTTCACCGCCACG  
CTCCATGCCGTGGGCTTCGTGCTGCCGCTGGCGGTGCTCTGCCCTCACCTCGCTCCAGGTGCAC  
CGGGTGGCACGCAGACACTGCCAGCGCATGGACACCGTCACCATGAAGGCGCTCGCGCTGCTC  
GCCGACCTGCACCCCAGTGTGCGGCAGCGCTGCCTCATCCAGCAGAAGCGGCGCCGCCACCGC  
GCCACCAGGAAGATTGGCATTGCTATTGCGACCTTCCTCATCTGCTTTGCCCCGTATGTCATG  
ACCAGGCTGGCGGAGCTCGTGCCCTTCGTCACCGTGAACGCCAGTGGGGCATCCTCAGCAAG  
TGCTGACCTACAGCAAGGCGGTGGCCGACCCGTTACGTA CTCTGCTCCGCCGGCCGTTTC  
CGCCAAGTCCTGGCCGGCATGGTGCACCGGCTGCTGAAGAGAACCCCGCGCCAGCATCCACC  
CATGACAGCTCTCTGGATGTGGCCGGCATGGTGCACCAGCTGCTGAAGAGAACCCCGCGCCCA  
GCGTCCACCCACAACGGCTCTGTGGACACAGAGAATGATTCTGCTGCCTGCAGCAGACACACT**TGA**  
GGGCTTGGCAGGGCTCATCGCCCCACCTTCTAAGAAGCCCTGTGGAAAGGGCACTGGCCCTG  
CCACAGAGATGCCACTGGGGACCCCCAGACACCAGTGGCTTGACTTTGAGCTAAGGCTGAG

595/615

**FIGURE 590**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA149930
><subunit 1 of 1, 363 aa, 1 stop
><MW: 39332, pI: 10.42, NX(S/T): 3
MGPGEALLAGLLVMVLAVALLSNALVLLCCAYSaelRTRASGVLLVNLSLGHLLLAALDM
PFTLLGVMRGRTPSAPGACQVIGFLDTFLASNAALSVAALSADQWLAVGFPLRYAGRRLRP
RYAGLLLGCAWGQSLAFSGAALGCSWLGYSSAFASCSLRLPPEPERPRFAAFTATLHAVG
FVLPLAVLCLTSLQVHRVARRHCQRMDTVMTKALALLADLHPSVRQRCLIQKRRRRHRAT
RKIGIAIATFLICFAPYVMTRLAELVPFVTVNAQWGILSKCLTYSKAVADPFTYSLLRRP
FRQVLAGMVHRLKRTPRPASTHDSsLDVAGMVHQLLKRTPRPASTHNGSVDtENDSCLQ
QTH
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-24

**Transmembrane domains:**

Amino acids 46-60;85-103;130-150;175-192;246-264

**N-glycosylation sites:**

Amino acids 47-51;348-352;355-359

**Tyrosine kinase phosphorylation site:**

Amino acids 286-295

**N-myristoylation sites:**Amino acids 66-72;124-130;128-134;132-138;139-145;244-250;  
349-355**G-protein coupled receptor proteins:**

Amino acids 72-112

**7 transmembrane receptor (rhodopsin family):**

Amino acids 22-294

596/615

**FIGURE 591**

AAC**ATG**GCTGCGGCGCCTGGGCTGCTCGTCTGGCTGCTCGTGCTCCGGCTGCCCTGGCGGGTG  
CCGGGCCAGCTGGACCCCAGCACTGGCCGGCGGTTCTCGGAGCACAACTCTGCGCGGACGAC  
GAATGCAGCATGATGTACCGCGGTGAGGCTCTTGAAGATTTACAGGCCCGGATTGTCGTTTT  
GTGAATTTTAAAAAAGGTGATCCTGTATATGTTTACTATAAACTGGCAAGAGGATGGCCTGAA  
GTTTGGGCTGGAAGTGTTGGACGCACTTTTGGATATTTTCCAAAAGATTTAATCCAGGTAGTT  
CATGAATATACCAAAGAAGAGCTACAAGTTCCAACAGATGAGACGGATTTTGTGTTGTTTTGAT  
GGAGGAAGAGATGATTTTCATAATTATAATGTAGAAGAACTTTTAGGGTTTTTGGAACTGTAC  
AATTCTGCAGCTACAGATTCTGAGAAAGCTGTAGAAAAAACTTTACAGGATATGGAAAAAAC  
CCTGAATTATCTAAGGAAAGGGAACCTGAACCTGAACCAGTAGAAGCCAACTCAGAGGAAAGT  
GATAGTGTATTCTCAGAAAACACTGAGGATCTTCAGGAACAGTTTACAACCTCAGAAGCACCAC  
TCCCATGCAAACAGCCAAGCAAATCATGCTCAGGGAGAGCAGGCTTCATTTGAATCTTTTGAA  
GAAATGCTGCAAGATAAACTAAAAGTGCCAGAAAGTGAAAACAACAAAACCAGCAATAGTTCT  
CAGGTCTCAAATGAACAGGATAAGATTGATGCCTATAAACTTTTGAAAAAAGAAATGACTCTA  
GACTTGAAAACCAAATTTGGCTCAACAGCTGATGCACTTGTATCTGATGATGAGACAACCAGA  
CTCGTTACTTCATTAGAAGATGATTTTGATGAGGAATTGGATACTGAGTATTATGCAGTTGGA  
AAGGAAGATGAGGAGAACCAAGAAGACTTTGATGAGTTGCCATTACTTACCTTTACAGATGGG  
GAAGATATGAAAACCTCAGCAAAGTCTGGCGTTGAGAAATATCCAACAGATAAAGAGCAGAAT  
TCAAATGAAGAGGACAAGGTTCACTAACTGTGCCCCCTGGCATCAAAAATGATGATAAAAAT  
ATACTAACAACCTGGGGGGACACTATCTTCTCTATTGTACAGGAGGTGAAGAAACAAGAGAT  
ACGATGGATTTAGAGAGCTCTAGTTCAGAGGAAGAAAAAGAAGATGATGATGATGCATTAGTC  
CCAGATAGCAAACAGGGGAAACCACAGTCAGCAACAGATTATAGTGACCCTGACAATGTAGAT  
GATGGTCTTTTTATTGTAGACATTCCTAAAACAAATAATGACAAAAGAAGTAAACGCAGAACAT  
CACATTAAAGGAAAAAGGGAGGGGAGTTCAAGGAATCCAAGAGGGGCCTGGTACAAGATGAGACA  
GAATTAGAGGATGAAAATCAAGAAGGCTTTAAAACAGAGCCATAAACTAT**TGA**CCTCTGAGG  
TTTCATTGGAAAGAAAGTGACTGTGCATTATCCATTACAGTAAAGGATTTCAATTGGCTTCAA  
AATCCAAAAGTTTATTTTAAAAGGTTTGTGTTAGAACTAAGCTGCCTTGGCAGTGTGCATTT  
TTGAGCCAAACAATTCAAAAATGTCATTTCTTCCCTAAATAAAAATCACCTTTTAAGCTAGAG  
CGTCCTTACAACCTTTGAAATGTGCAATAAAGAATACCTGTGTTTTAGCTAATGTAGCATATGT  
AATTGCAAAATGATTTAGAATGTCATGAAAAATATGAACATTTCTGTGGAAATGCTTTAAGA  
ACATGTATTTCCATTATCCTATTTTGTAGTGTACACCAGCTGAATACGGAGCAATGGTGTGTTAT  
AAGCGTTTTTTTTAACTATCTGGTCACAAAGACTGTTACGCTAAAAATGTTTACTAAAAGATC  
ACTAACTATCTCCCCTCTTGCTGAAGTTCTTTGTAGTAATAGCTCATAAAAATTTGTTTATT  
AATATTTAAAAA

597/615

**FIGURE 592**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA150157
><subunit 1 of 1, 499 aa, 1 stop
><MW: 56471, pI: 4.31, NX(S/T): 2
MAAAPGLLVWLLVLRRLPWRVPGQLDPSTGRRFSEHKLCADDECSSMYRGEALEDFTGPDC
RFVNFKKGDPVYVYYKLARGWPEVWAGSVGRTFGYFPKDLIQVVHEYTKHEELQVPTDET
FVCFDGGRRDDFHNYNVEELLGFLELYNSAATDSEKAVEKTLQDMEKNPELSKEREPEPEP
VEANSEESDSVFSENTEDLQEQFTTQKHHSHANSQANHAQGEQASFESEMLQDKLKVP
ESENKTSNSSQVSNEQDKIDAYKLLKKEMTLDLKTGFGSTADALVSDDDETTRLVTSLED
DFDEELDT EYYAVGKEDEENQEDFDELPLLTFTDGEDMKTPAKSGVEKYPTDKEQNSNEE
DKVQLTVPPGIKNDDKNILTTWGDITFSIVTGGEETRDMDLESSSSEEEKEDDDALVP
DSKQKGKQSATDYSDDPNVDDGLFIVDIPKTNNKENVNAEHHIKGKGRGVQESKRGLVQD
ETELEDENQEGFKTEPIKL
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-22

**N-glycosylation sites:**

Amino acids 245-249;249-253

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 30-34

**Tyrosine kinase phosphorylation site:**

Amino acids 66-72

**N-myristoylation sites:**

Amino acids 392-398;469-475

**Amidation site:**

Amino acids 28-32

**Aminoacyl-transfer RNA synthetases class-II signature 1:**

Amino acids 47-70



598/615

**FIGURE 593**

GGGCCAGTAGAGTGTGTCTGGGTCAGCTGAGTGACTACATCAAAGCTCCCAGCCTTGAAAAAC  
ACATGCTGTTCCCAGGCCTCAAGATATTGAAACATTAATTAGATAATTTAAAGTAGCGTTTTTC  
TTCTACA**ATGT**CTGAAGAAGTGACCTACGCGACACTCACATTTCAGGATTCTGCTGGAGCAAG  
GAATAACCGAGATGGAAATAACCTAAGAAAAAGAGGGCATCCAGCTCCATCTCCCATTTGGCG  
TCATGCTGCTCTGGGTCCTGGTAACCTCTTTCCTGATGTTGCTGATTGGGCTGGTGACGTTGGG  
GATGATGTTTTTGCAGATATCTAATGACATTAACCTCAGATTCAGAGAAATTGAGTCAACTTCA  
GAAAACCATCCAACAGCAGCAGGATAACTTATCCCAGCAACTGGGCAACTCCAACAACCTTGTC  
CATGGAGGAGGAATTTCTCAAGTCACAGATCTCCAGTCTACTGAAGAGGCAGGAACAAATGGC  
CATCAAACCTGTGCCAAGAGCTAATCATTACATACTTCAGACCACAGATGTAATCCATGTCCTAA  
GATGTGGCAATGGTACCAAAATAGTTGCTACTATTTTACAACAAATGAGGAGAAAACCTGGGC  
TAACAGTAGAAAAGGACTGCATAGACAAGAACTCCACCCTAGTGAAGATAGACAGTTTGGAAGA  
AAAGGATTTTCTTATGTCACAGCCATTACTCATGTTTTTCGTTCTTTTGGCTGGGATTATCATG  
GGACTCCTCTGGCAGAAGTTGGTCTGGGAAGATGGCTCTGTTCCCTCTCCATCCTTGTACGT  
CTCTAACTAT**TGA**GGGTAAACACAAGCTTTCCATGGAATCCTGGGAAAATTAATAATGATTGT  
GAGAATTATAAATACAGACATAAAAAGAGGAGTACAACATACTGAGAAAAGAGCTCCAGTAAC  
AAATATTGAAAGGAGATTTAGTACTAAAGAACTTGACCAGATCAATGGATCCAAGGATGTGC  
TTATTTTCAAAAAGGAAATATTTATATTTCTCGCTGTAGTGCTGAAATTTTTTGGATTGCGA  
GAAGACAGCTGCCCCAGTGAAGACTGAGGATTTGGATTAGTATGCTTCTTCCAAATTCTCCAA  
GAAGTAAGAGACTTGTGAGTAAGCTCATATGAGGAAAGAGGAAACTACGGTACCAGAGCAAGG  
CGAATTCTGCA

599/615

**FIGURE 594**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA150163
><subunit 1 of 1, 232 aa, 1 stop
><MW: 26754, pI: 5.80, NX(S/T): 3
MSEEVTYATLTFQDSAGARNNRDGNLRLKRGHPAPSPIWRHAALGLVTLCLMLLIGLVTL
GMMFLQISNDINSDEKLSQLQKTIQQQQDNLSQLGNSNNLSMEEEFLLKSQISSLLKRQ
EQMAIKLCQELIIHTSDHRCNCPKMWQWYQNSCYFTTNEEKTWANSRKDCIDKNSTLV
KIDSLEEKDFLMSQPLLMFSFFWLGLSWDSSGRSWFWEDGSVPSPSLYVSNY
```

**Important features of the protein:****Transmembrane domain:**

Amino acids 42-62

**N-glycosylation sites:**

Amino acids 91-95;101-105;176-180

**N-myristoylation sites:**

Amino acids 17-23;97-103

600/615

**FIGURE 595**

CGGACGCGTGGGGAAGAGGAGGAGGAGGAAGAAGACGTGGACAAGGACCCCCATCCTACCCAG  
AACACCTGCCTGCGCTGCCGCCACTTCTCTTTAAGGGAGAGGAAAAGAGAGCCTAGGAGAACC  
**ATG**GGGGGCTGCGAAGTCCGGGAATTTCTTTTGCAATTTGGTTTCTTCTTGCCATATGCTGACA  
GCGTGGCCAGGCGACTGCAGTCACGTCTCCAACAACCAAGTTGTGTTGCTTGATACAACAAC  
GTACTGGGAGAGCTAGGATGGAAAACATATCCATTAAATGGGTGGGATGCCATCACTGAAATG  
GATGAACATAATAGGCCCATTCACACATACCAGGTATGTAATGTAATGGAACCAAACCAAAAC  
AACTGGCTTCGTACAAACTGGATCTCCCGTGATGCAGCTCAGAAAATTTATGTGGAAATGAAA  
TTCACACTAAGGGATTGTAACAGCATCCCATGGGTCTTGGGGACTTGCAAAGAAACATTTAAT  
CTGTTTTATATGGAATCAGATGAGTCCACGGAATTAAATTCAAGCCAAACCAGTATACAAAG  
ATCGACACAATTGCTGCTGATGAGAGTTTTACCCAGATGGATTGGGTGATCGCATCCTCAAA  
CTCAACACTGAAATTCGTGAGGTGGGGCCTATAGAAAGGAAAGGATTTTATCTGGCTTTTCAA  
GACATTGGGGCGTGCAATTGCCCTGGTTTCAGTCCGTGTTTTCTACAAGAAATGCCCTTCACT  
GTTTCGTAAC TTGGCCATGTTTCCTGATACCATTCCAAGGGTTGATTCCCTCCTCTTTGGTTGAA  
GTACGGGGTTCTTGTGTGAAGAGTGCTGAAGAGCGTGACACTCCTAACTGTATTGTGGAGCT  
GATGGAGATTGGCTGGTTCCTCTTGGAAGGTGCATCTGCAGTACAGGATATGAAGAAATTGAG  
GGTTCTTGCCATGGAGCCTCCAAAGGCCGCTGCTTCT**TAG**TTGGCCATCTTGGCCCCACCCGA  
AACAGTAACCTTTGAAGAATAAAAAGAAAAAGCAAAAGAGTAGCATTACTAAAATATTAAACGG  
TTACATTTACAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

601/615

**FIGURE 596**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA153579
><subunit 1 of 1, 285 aa, 1 stop
><MW: 32368, pI: 5.32, NX(S/T): 0
MGGCEVREFLLQFGFFLPMLTAWPGDCSHVSNNQVVLLDTTTVLGELGWKTYPLNGWDAI
TEMDEHNRPIHTYQVCNVMEPNQNNWLRTNWI SRDAAQKIYVEMKFTLRDCNSIPWVLGT
CKETFNLFYMESDESHGIKFKPNQYTKIDTIAADESFTQMDLGDRILKLNTEIREVGPIE
RKGFYLAFAQDIGACIALVSVRVFYKKCPFTVRNLAMFPDTIPRVDSSSLVEVRGSCVKSA
EERDTPKLYCGADGDWLVPLGRCICSTGYEEIEGSCHGASKGRCF
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-22

**N-myristoylation sites:**

Amino acids 192-198;274-280;278-284

**Receptor tyrosine kinase class V signature 1:**

Amino acids 192-209

**Ephrin receptor ligand binding domain:**

Amino acids 34-207

602/615

**FIGURE 597**

ACACTGGCCAAACAAAAACGAAAGCACTCCGTGCTGGAAGTAGGAGGAGAGTCAGGACTCCCA  
GGACAGAGAGTGCACAACTACCCAGCACAGCCCCCTCCGCCCCCTCTGGAGGCTGAAGAGGG  
ATTCCAGCCCCTGCCACCCACAGACACGGGCTGACTGGGGTGTCTGCCCCCTTGGGGGGGG  
CAGCACAGGGCCTCAGGCCTGGGTGCCACCTGGCACCTAGAAG**ATG**CCTGTGCCCTGGTTCTT  
GCTGTCCTTGGCACTGGGCCGAAGCCCAGTGGTCCTTTCTCTGGAGAGGCTTGTGGGGCCTCA  
GGACGCTACCCACTGCTCTCCGGGCCTCTCCTGCCGCCTCTGGGACAGTGACATACTCTGCCT  
GCCTGGGGACATCGTGCTGCTCCGGGGCCCCGTGCTGGCGCCTACGCACCTGCAGACAGAGCT  
GGTGCTGAGGTGCCAGAAGGAGACCGACTGTGACCTCTGTCTGCGTGTGGCTGTCCACTTGGC  
CGTGCAATGGGCACTGGGAAGAGCCTGAAGATGAGGAAAAGTTTGGAGGAGCAGCTGACTCAGG  
GGTGGAGGAGCCTAGGAATGCCTCTCTCCAGGCCCAAGTCGTGCTCTCCTTCCAGGCCTACCC  
TACTGCCCCGCTGCGTCCTGCTGGAGGTGCAAGTGCCTGCTGCCCTTGTGCAGTTTGGTCAGTC  
TGTGGGCTCTGTGGTATATGACTGCTTCGAGGCTGCCCTAGGGAGTGAGGTACGAATCTGGTC  
CTATACTCAGCCCAGGTACGAGAAGGAACTCAACCACACACAGCAGCTGCCTGCCCTGCCCTG  
GCTCAACGTGTGACGAGATGGTGACAACGTGCATCTGGTTCTGAATGTCTCTGAGGAGCAGCA  
CTTCGGCCTCTCCCTGTACTGGAATCAGGTCCAGGGCCCCCCCCAAAACCCCGGTGGCACAAAA  
CCTGACTGGACCGCAGATCATTACCTTGAACCACACAGACCTGGTTCCCTGCCTCTGTATTCA  
GGTGTGGCCTCTGGAACCTGACTCCGTTAGGACGAACATCTGCCCTTCAGGGAGGACCCCCG  
CGCACACCAGAACCTCTGGCAAGCCGCCGACTGCGACTGCTGACCCTGCAGAGCTGGCTGCT  
GGACGCACCCTGCTCGCTGCCCGCAGAAGCGGCACTGTGCTGGCGGGCTCCGGGTGGGGACCC  
CTGCCAGCCACTGGTCCCACCGCTTTCTGAGGAGAACGTCACTGTGGACAAGGTTCTCGAGTT  
CCCATTGCTGAAAGGCCACCCTAACCTCTGTGTTTCAGGTGAACAGCTCGGAGAAGCTGCAGCT  
GCAGGAGTGCTTGTGGGCTGACTCCCTGGGGCCTCTCAAAGACGATGTGCTACTGTTGGAGAC  
ACGAGGCCCCCAGGACAACAGATCCCTCTGTGCCTTGGAAACCAGTGGCTGTACTTCACTACC  
CAGCAAAGCCTCCACGAGGGCAGCTCGCCTTGGAGAGTACTTACTACAAGACCTGCAGTCAGG  
CCAGTGTCTGCAGCTATGGGACGATGACTTGGGAGCGCTATGGGCCTGCCCCATGGACAAATA  
CATCCACAAGCGCTGGGCCCTCGTGTGGCTGGCCTGCCTACTCTTTGCCGCTGCGCTTTCCCT  
CATCCTCCTTCTCAAAAAGGATCACGCGAAAGGGTGGCTGAGGCTCTTGAAACAGGACGTCCG  
CTCGGGGGCGGCCGCCAGGGGCCGCGCGGCTCTGCTCCTCTACTCAGCCGATGACTCGGGTTT  
CGAGCGCCTGGTGGGCGCCCTGGCGTCCGGCCCTGTGCCAGCTGCCGCTGCGCGTGGCCGTAGA  
CCTGTGGAGCCGTCGTGAACCTGAGCGCGCAGGGGGCCCGTGGCTTGGTTTCACGCGCAGCGCG  
CCAGACCCTGCAGGAGGGCGGCGTGGTGGTCTTGCTCTTCTCTCCGGTGCGGTGGCGCTGTG  
CAGCGAGTGGCTACAGGATGGGGTGTCCGGGCCCGGGGCGCACGGCCCGCACGACGCCTTCCG  
CGCCTCGCTCAGCTGCGTGCTGCCGACTTCTTGACAGGGCCGGGCGCCCGGAGCTACGTGGG  
GGCCTGCTTCGACAGGCTGCTCCACCCGACGCGGTACCCGCCCTTTTCCGCACCGTGCCCGT  
CTTCACACTGCCCTCCCAACTGCCAGACTTCTGGGGGCCCTGCAGCAGCCCTCGCGCCCCGCG  
TTCCGGGCGGCTCCAAGAGAGAGCGGAGCAAGTGTCCCGGGCCCTTCAGCCAGCCCTGGATAG  
CTACTTCCATCCCCCGGGGACTCCCGCGCCGGGACGCGGGGTGGGACCAGGGGCGGGACCTGG  
GGCGGGGGACGGGACT**TAA**ATAAAGGCAGACGCTGTTTTCTAAAAAA

603/615

**FIGURE 598**

```
></usr/seqdb2/sst/DNA/Dnaseqs.full/ss.DNA164625
><subunit 1 of 1, 705 aa, 1 stop
/><MW: 76970, pI: 6.00, NX(S/T): 9
MPVPWFLLSLALGRSPVVLSELRVGPQDATHCSPGLSCRLWDSILCLPGDIVPAPGP
VLAPTHLQTELVLRCQKETDCDLCLRVAVHLAVHGHWEPEDEEKFGGAADSGVEEPRN
ASLQAQVVLSEFQAYPTARCVLLEVQVPAALVQFGQSVGSVVYDCFEAALGSEVRIWSYT
QPRYEKELNHTQQLPALPWLNVSADGDNVHLVLNVSEEQHFGLSLYWNQVQGPPKPRWH
KNLTGPGQIITLNHTDLVPCLCIQVWPLEPDSVRTNICPFREDPRAHQNLWQAARLRLLT
LQSWLLDAPCSLPAEAALCWRAPGGDPCQPLVPPLSWENVTVDKVLEFPLLKGHPNLCV
QVNSSEKLQEQELWADSLGPLKDDVLLLETRGPDNRSLCALEPSGCTSLPSKASTRA
ARLGEYLLQDLQSGQCLQLWDDDLGALWACPMCKYIHKRWALVWLACLLFAAALSLILL
LKKDHAKGWLRLLLKQDVRSAAAARGRAALLLYSADDSGFERLVGALASALCQLPLRVAV
DLWSRRELSAQGPVAVFHAQRRQTLQEGGVVLLLFSPGAVALCSEWLQDGVSGPGAHP
HDAFRASLSCVLPDFLQGRAPGSYVGACFDRLPHDAVPALFRTVPVFTLPSQLPDFLG
ALQQPRAPRSGRLQERAEQVSRALQPALDSYFHPPGTPAPGRGVGPGAGPGAGDGT
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-20

**Transmembrane domain:**

Amino acids 453-473

**N-glycosylation sites:**Amino acids 118-122;186-190;198-202;211-215;238-242;  
248-252;334-338;357-360;391-395;**Glycosaminoglycan attachment site:**

Amino acids 583-587

**cAMP- and cGMP-dependent protein kinase phosphorylation site:**

Amino acids 552-556

**N-myristoylation sites:**Amino acids 107-113;152-158;319-325;438-444;516-522;612-618;  
692-698;696-702;700-706

604/615

**FIGURE 599**

GGTCCTTA**ATG**GCAGCAGCCGCCGCTACCAAGATCCTTCTGTGCCTCCCGCTTCTGCTCCTGC  
TGTCGGCTGGTCCCGGGCTGGGCGAGCCGACCCTCACTCTCTTTGCTATGACATCACCGTCA  
TCCCTAAGTTTCAACCTGGACCACGGTGGTGTGCGGTTCAAGGCCAGGTGGATGAAAAGACTT  
TTCTTCACTATGACTGTGGCAACAAGACAGTCACACCTGTCAGTCCCCTGGGGAAGAACTAA  
ATGTCACAACGGCCTGGAAAGCACAGAACCCAGTACTGAGAGAGGTGGTGGACATACTTACAG  
AGCAACTGCGTGACATTGAGCTGGAGAATTACACACCCAAGGAACCCCTCACCTGCAGGCAA  
GGATGTCTTGTGAGCAGAAAGCTGAAGGACACAGCAGTGGATCTTGGCAGTTCAGTTTCGATG  
GGCAGATCTTCCTCCTCTTTGACTCAGAGAAGAGAATGTGGACAACGGTTCATCCTGGAGCCA  
GAAAGATGAAAGAAAAGTGGGAGAATGACAAGGTTGTGGCCATGTCCTTCCATTACTTCTCAA  
TGGGAGACTGTATAGGATGGCTTGAGGACTTCTTGATGGGCATGGACAGCACCCCTGGAGCCAA  
GTGCAGGAGCACCACTCGCCATGTCCTCAGGCACAACCCAACCTCAGGGCCACAGCCACCACCC  
TCATCCTTTGCTGCCTCCTCATCATCCTCCCCTGCTTCATCCTCCCTGGCATCT**GAG**AGAGAGT  
CCTTTAGAGTGACAGGTTAAAGCTGATACCAAAGGCTCCTGTGAGCACGGTCTTGATCAAAC  
TCGCCCTTCTGTCTGGCCAGCTGCCACGACCTACGGTGTATGTCCAGTGGCCTCCAGCAGAT  
CATGATGACATCATGGACCAATAGCTCATTCAGTGCCTTGATTCCTTTTGCCAACAATTTTA  
CCAGCAGTTATACCTAACATATTATGCAATTTTCTCTTGGTGCTACCTGATGGAATTCCTGCA  
CTTAAAGTTCTGGCTGACTAAACAAGATATATCATTTTCTTTCTTCTTTTGTGTTGGAAAA  
TCAAGTACTTCTTTGAATGATGATCTCTTTCTTGCAAATGATATTGTCAGTAAAATAATCACG  
TTAGACTTCAGACCTCTGGGGATTCTTTCCGTGTCCTGAAAGAGAATTTTAAATTATTTAAT  
AAGAAAAAATTTATATTAATGATTGTTTCCTTTAGTAATTTATTGTTCTGTACTGATATTTAA  
ATAAAGAGTTCTATTTCCCAAAAAAAAAAAAAAAAAAAAAA

605/615

**FIGURE 600**

MAAAAATKILLCLPLLLLLSGWSRAGRADPHSLCYDITVIPKFRPGPRWCAVQGQVDEKTFLH  
YDCGNKTVTPVSPPLGKKLVTTAWKAQNPVLREVVDILTEQLRDIQLENYTPKEPLTLQARMS  
CEQKAEGHSSGSWQFSFDGQIFLLFDSEKRMWTTVHPGARKMKEKWENDKVVAMSFHYFSMGD  
CIGWLEDFLMGMDSTLEPSAGAPLAMSSGTTQLRATATTLILCCLLIILPCFILPGI

**Important features:****Signal peptide:**

amino acids 1-25

**Transmembrane domain:**

amino acids 224-246

**N-glycosylation site.**

amino acids 68-72, 82-86

**N-myristoylation site.**

amino acids 200-206, 210-216

**Amidation site.**

amino acids 77-81



606/615

**FIGURE 601**

GCAGTCAGAGACTTCCCCTGCCCTCGCTGGGAAAGAACATTAGGAATGCCTTTTAGTGCCTT  
GCTTCCTGAAGTAGCTCACAGTAGCCCGCGGCCAGGGCAATCCGACCACATTTCACTCTCA  
CCGCTGTAGGAATCCAG**ATG**CAGGCCAAGTACAGCAGCACGAGGGACATGCTGGATGATGATG  
GGGACACCACCATGAGCCTGCATTCTCAAGCCTCTGCCACAACCTCGGCATCCAGAGCCCCGGC  
GCACAGAGCACAGGGCTCCCTCTTCAACGTGGCGACCAGTGGCCCTGACCCTGCTGACTTTGT  
GCTTGGTGCTGCTGATAGGGCTGGCAGCCCTGGGGCTTTTGTTTTTTTCAGTACTACCAGCTCT  
CCAATACTGGTCAAGACACCATTTCTCAAATGGAAGAAAGATTAGGAAATACGTCCCAAGAGT  
TGCAATCTCTTCAAGTCCAGAATATAAAGCTTGCAGGAAGTCTGCAGCATGTGGCTGAAAAAC  
TCTGTCGTGAGCTGTATAACAAAGCTGGAGCACACAGGTGCAGCCCTTGTACAGAACAATGGA  
AATGGCATGGAGACAATTGCTACCAGTTCTATAAAGACAGCAAAAGTTGGGAGGACTGTAAAT  
ATTTCTGCCTTAGTGAAAACCTACCATGCTGAAGATAAACAAACAAGAAGACCTGGAATTTG  
CCGCGTCTCAGAGCTACTCTGAGTTTTTCTACTCTTATTGGACAGGGCTTTTGCGCCCTGACA  
GTGGCAAGGCCTGGCTGTGGATGGATGGAACCCCTTTCACTTCTGAACTGTTCCATATTATAA  
TAGATGTCACCAGCCCAAGAAGCAGAGACTGTGTGGCCATCCTCAATGGGATGATCTTCTCAA  
AGGACTGCAAAGAATTGAAGCGTTGTGTCTGTGAGAGAAGGGCAGGAATGGTGAAGCCAGAGA  
GCCTCCATGTCCCCCTGAAACATTAGGCGAAGGTGACT**TGA**TTCGCCCTCTGCAACTACAAAT  
AGCAGAGTGAGCCAGGCGGTGCCAAGCAAGGGCTAGTTGAGACATTGGGAAATGGAACATAA  
TCAGGAAAGACTATCTCTCTGACTAGTACAAAATGGGTTCCTCGTGTTTCCTGTTTCAGGATCAC  
CAGCATTTCTGAGCTTGGGTTTATGCACGTATTTAACAGTCACAAGAAGTCTTATTTACATGC  
CACCAACCAACCTCAGAAACCCATAATGTCATCTGCCTTCTTGGCTTAGAGATAACTTTTAGC  
TCTCTTTCTTCTCAATGTCTAATATCACCTCCCTGTTTTTCATGTCTTCCTTACACTTGGTGGA  
ATAAGAAACTTTTTGAAGTAGAGGAAATACATTGAGGTAACATCCTTTTCTCTGACAGTCAAG  
TAGTCCATCAGAAATTGGCAGTCACTTCCCAGATTGTACCAGCAAATACACAAGGAATTCTTT  
TTGTTTGTTCAGTTCATACTAGTCCCTTCCCAATCCATCAGTAAAGACCCCATCTGCCTTGT  
CCATGCCGTTTCCCAACAGGGATGTCACCTTGATATGAGAATCTCAAATCTCAATGCCTTATAA  
GCATTCCTTCCTGTGTCCATTAAGACTCTGATAATTGTCTCCCCTCCATAGGAATTTCTCCCA  
GGAAAGAAATATATCCCATCTCCGTTTCATATCAGAACTACCGTCCCCGATATTCCCTTCAG  
AGAGATTAAAGACCAGAAAAAAGTGAGCCTCTTCATCTGCACCTGTAATAGTTTCAGTTCCTA  
TTTTCTTCATTGACCCATATTTATACCTTTCAGGTACTGAAGATTTAATAATAATAAATGTA  
AATACTGTGAAAAA

607/615

**FIGURE 602**

MQAKYSSTRDMLDDDGDTTMSLHSQASATTRHPEPRRTEHRAPSSTWRPVALTLLTLCVLVLLI  
GLAALGLLFFQYYQLSNTGQDTISQMEERLGNTSQELQSLQVQNIKLAGSLQHVAEKLCRELY  
NKAGAHRCSPCTEQWKWHGDNCYQFYKDSKSWEDCKYFCLSENSTMLKINKQEDLEFAASQSY  
SEFFYSYWTGLLRPDSGKAWLWMDGTPFTSELFHIIIDVTSPRSRDCVAILNGMIFSKDCKEL  
KRCVCERRAGMVKPESLHVPPETLGED

608/615

**FIGURE 603**

GGGAGAGAGGATAAATAGCAGCGTGGCTTCCCTGGCTCCTCTCTGCATCCTTCCCGACCTTCC  
CAGCAAT**ATG**CATCTTGCACGTCTGGTCGGCTCCTGCTCCCTCCTTCTGCTACTGGGGGCCCT  
GTCTGGATGGGCGGCCAGCGATGACCCCAATTGAGAAGGTCATTGAAGGGATCAACCGAGGGCT  
GAGCAATGCAGAGAGAGAGAGGTGGGCAAGGCCCTGGATGGCATCAACAGTGAATCACGCATGC  
CGGAAGGGAAGTGGAGAAGGTTTTCAACGGACTTAGCAACATGGGGAGCCACACCGGCAAGGA  
GTTGGACAAAGGCGTCCAGGGGCTCAACCACGGCATGGACAAGGTTGCCCATGAGATCAACCA  
TGGTATTGGACAAGCAGGAAAGGAAGCAGAGAAGCTTGGCCATGGGGTCAACAACGCTGCTGG  
ACAGGCCGGGAAGGAAGCAGACAAAGCGGTCCAAGGGTTCCACACTGGGGTCCACCAGGCTGG  
GAAGGAAGCAGAGAACTTGGCCAAGGGGTCAACCATGCTGCTGACCAGGCTGGAAAGGAAGT  
GGAGAAGCTTGGCCAAGGTGCCCAACCATGCTGCTGGCCAGGCCGGAAGGAGCTGCAGAATGC  
TCATAATGGGGTCAACCAAGCCAGCAAGGAGGCCAACCAGCTGCTGAATGGCAACCATCAAAG  
CGGATCTTCCAGCCATCAAGGAGGGGCCACAACCACGCCGTTAGCCTCTGGGGCCTCAGTCAA  
CACGCCTTTCATCAACCTTCCCGCCCTGTGGAGGAGCGTCGCCAACATCATGCCCT**TAA**ACTGG  
CATCCGGCCTTGCTGGGAGAATAATGTCGCCGTTGTCACATCAGCTGACATGACCTGGAGGGG  
TTGGGGGTGGGGGACAGGTTTCTGAAATCCCTGAAGGGGGTTGTAAGGGGTTGTAATAAA  
CTTGATACACCA

609/615

**FIGURE 604**

></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA66675  
><subunit 1 of 1, 247 aa, 1 stop  
><MW: 25335, pI: 7.00, NX(S/T): 0  
MHLARLVGSCSLLLLLLGALSGWAASDDPIEKVIEGINRGLSNAEREVKGALDGINSGITHAGR  
EVEKVFNGLSNMGSHGTGKELDKGVQGLNHGMDKVAHEINHGIGQAGKEAEKLGHG VNNAAGQA  
GKEADKAVQGFHTGVHQAGKEAEKLGQGVNHAADQAGKEVEKLGQGAHHAAGQAGKELQNAHN  
GVNQASKEANQLLNGNHQSGSSSHQGGATTTPLASGASVNTPFINLPALWRSVANIMP

**Important features of the protein:**

**Signal peptide:**

amino acids 1-25

**Homologous region to circumsporozoite (CS) repeats:**

amino acids 35-225

610/615

**FIGURE 605**

GCGACGCGCGGCGGGGCGGCGAGAGGAAACGCGGCGCCGGGCGGGCCCGGCCCTGGAG**ATGG**  
TCCCCGCGCGCCGCGGGGCTGGTGTTGTCTCGTGCTCTGGCTCCCCGCGTGCGTTCGCGGCCACG  
GCTTCCGTATCCATGATTATTTGTACTTTCAAGTGCTGAGTCCTGGGGACATTCGATACATCT  
TCACAGCCACACCTGCCAAGGACTTTGGTGGTATCTTTACACAAGGTATGAGCAGATTACACC  
TTGTCCCCGCTGAACCTCCAGAGGCCTGCGGGGAACCTCAGCAACGGTTTCTTCATCCAGGACC  
AGATTGCTCTGGTGGAGAGGGGGGGGCTGCTCCTTCCTCTCCAAGACTCGGGTGGTCCAGGAGC  
ACGGCGGGCGGGCGGTGATCATCTCTGACAACGCAGTTGACAATGACAGCTTCTACGTGGAGA  
TGATCCAGGACAGTACCCAGCGCACAGCTGACATCCCCGCCCTCTTCCTGCTCGGCCGAGACG  
GCTACATGATCCGCCGCTCTCTGGAACAGCATGGGCTGCCATGGGCCATCATTTCCATCCCAG  
TCAATGTCACCAGCATCCCCACCTTTGAGCTGCTGCAACCGCCCTGGACCTTCTGG**TAGAAGA**  
GTTTGTCCACATTCCAGCCATAAGTGACTCTGAGCTGGGAAGGGGAAACCCAGGAATTTTGC  
TACTTGGAATTTGGAGATAGCATCTGGGGACAAGTGGAGCCAGGTAGAGGAAAAGGGTTTGGG  
CGTTGCTAGGCTGAAAGGGAAGCCACCACTGGCCTTCCCTTCCCCAGGGCCCCCAAGGGTG  
TCTCATGCTACAAGAAGAGGCAAGAGACAGGCCCCAGGGCTTCTGGCTAGAACCCGAAACAAA  
AGGAGCTGAAGGCAGGTGGCCTGAGAGCCATCTGTGACCTGTCACACTCACCTGGCTCCAGCC  
TCCCCTACCCAGGGTCTCTGCACAGTGACCTTCACAGCAGTTGTTGGAGTGGTTTAAAGAGCT  
GGTGTGTTGGGGACTCAATAAACCTCACTGACTTTTGTAGCAATAAAGCTTCTCATCAGGGTTG  
CAAAAAAAAAAAAAAAAAAAAAAAAAA

611/615

**FIGURE 606**

&gt;&lt;/usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA76532

&gt;&lt;subunit 1 of 1, 188 aa, 1 stop

&gt;&lt;MW: 21042, pI: 5.36, NX(S/T): 2

MVPGAAGWCCLVLWLPACVAAHGFRIHDYLYFQVLSPGDIRYIFTATPAKDFGGIFHTRYEQI  
HLVPAEPPEACGELSNGFFIQDQIALVERGGCSFLSKTRVVQEHGGRAVIISDनावदन्सFYV  
EMIQDSTQRTADIPALFLLGRDGYMIRRSLEQHGLPWAIISIPVNVTSIPTFELLQPPWTFW**Signal peptide:**

amino acids 1-20

612/615

**FIGURE 607**

GCATTTGCCACTGGTTGCAGATCAGGCGGACGAGGAGCCGGGAAGGCAGAGCC**ATGT**GGCTGC  
CCCCTGCTCTGCTCCTTCTCAGCCTCTCAGGCTGTTTCTCCATCCAAGGCCAGAGTCTGTGA  
GAGCCCCAGAGCAGGGGTCCCTGACGGTTCAATGCCACTATAAGCAAGGATGGGAGACCTACA  
TTAAGTGGTGGTGCCGAGGGGTGCGCTGGGATACATGCAAGATCCTCATTGAAACCAGAGGGT  
CGGAGCAAGGAGAGAAGAGTGACCGTGTGTCCATCAAGGACAATCAGAAAGACCGCACGTTCA  
CTGTGACCATGGAGGGGCTCAGGCGAGATGACGCAGATGTTTACTGGTGTGGGATTGAAAGAA  
GAGGACCTGACCTTGGGACTCAAGTGAAAGTGATCGTTGACCCAGAGGGAGCGGCTTCCACAA  
CAGCAAGCTCACCTACCAACAGCAATATGGCAGTGTTTCATCGGCTCCCACAAGAGGAACCACT  
ACATGCTCCTGGTATTTGTGAAGGTGCCATCTTGCTCATCTTGGTCACTGCCATCCTCTGGT  
TGAAGGGGTCTCAGAGGGTCCCTGAGGAGCCAGGGGAACAGCCTATCTACATGAACTTCTCCG  
AACCTCTGACTAAAGACATGGCCACT**TAG**AGAGATGGATCTGCAGAGCCTTCTGCCCTGGCC  
ACGTTTCCAGAAGAGACTCGGGCTGTGGAAGGAACATCTACGAGTCTCGGGATGCAGTGACT  
GAGATAGGGGCCCTGGGCCCTCCGCCCTGGCCTTGGAGCTGGTGGGCACCTCCCTGTTCTGCAC  
AGCTCAGGGACTTAGCCAGGTCTCTCCTGAGCCACCATCACCTCCTGGGGTGCCAGCACCTG  
TTCTCTTGGTCAGGAGCTGTAGAGATGGAGCTCAAGCACTGGACGACTCTGTCCCCACTGCTG  
GAATAACTCGGGCACAGAGCATGGGACCAAAGTACAGAAAGAGGTTGGGGGAGACCCCCCAG  
CCCTAGACTTCCATCATTCCGGAGACCAACTCAACACCGTCTTGCCTGAGAACCTGATATATCC  
GTGTTTTTAAATTTTTTTTTTTCTAGCAAAGTTGGGTTTTAATGACTTATGTTTCATAGGAAAC  
CTCTCTGATCCCACACACAAGGAGGGTGATTCTGGGATGAGTTCCTGGTTCTAGGGCATGAGG  
GGCTGGATGGACCCTGTCCCCAGGGAGGACATGGCTCTGAGTCCACAGGGCTGAGGAGGCAAT  
GGGAACCTCCCTGGCCCCGGCCCGGTGCTTGCTCCTCCCCCTCCACCTCTTCTCCTCCTAGCT  
CCCCAAGCTCCCTGCCTATTCCCCACCTCCGAGGGGCTGCAGCTTGGGAGCCTCCTCAGCAT  
GACAGCTTGGGTCTCCTCCCCAAAAGAGCCTGTCAGGCCTCAAGAACCACCTCCAGGTGGGGA  
GGGCAGTAACGAAAACCATCGCAGGAAATGGCACCCCTCCCTTTTCGGTGATGTTGAAATCATG  
TTACTAATGAAAACGTCTCCTAGGGAAGTGGTCTGTCTCCTCACAGGCTTCACCCACGGCGAT  
GAGGCCCTTGAATGTGGTCACTTTGTGCTGTATGGTTGAGGGACCCTCACACCAAAGGGACCT  
TCCCATGTGAGATGTGCTCCCGCCCCACCTGCCCACAAGCAAACACACCACACATGTTCCGGC  
ATGTTGCCCTTTGAACACCCATGAGGACGCCTCCAACCTGCTCTTGGTTCTAATAGGGAGTAC  
TGACTGTCAGCAGTGGATAAAGGAGAGGGGACCCTCTGGTCCCTAGCATGGCACCCAGAGCCT  
CCCCTCTTCTTGCTCCTTCAGCCAAAGAGAACTTTCTCTGACTTTGAACTGAATTTAGGTCTC  
TGGCCAATGATGGGCCTGAAAATTCATAATGGCCAGAGAGGAGAGTTTCGAGCCCCGGCTAAGA  
TCCCCTGAGTCATTCTGTGAGGGACCAAGACCCACAGTCCACCAGCCCCAGGGCCCTACCTCC  
TGGAATGCTTTCCTGGATCCAGCTTCCCGAAGATCCGACCAGACCCAGGGAGGACGGCACCGC  
TCCGCGGGAGGGAAAGCCAAAGCATGGTGCTTCACCAGCTGGACTCAGGGGCGAGGGGACATG  
GGCGCTTGTCAACGTGATGTCATTCTTTTCCACCGTTTCTTCTGTTGATATTCAATGAATC  
CGTCAATCTCTCTGGGAAA

613/615

**FIGURE 608**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA105849
><subunit 1 of 1, 201 aa, 1 stop
><MW: 22689, pI: 7.41, NX(S/T): 1
MWLPPALLLLSLSGCFSIQGPESVRAPEQGSITVQCHYKQGWETYIKWWCRGVRWDTCKI
LIETRGSEQGEKSDRVSIKDNQKDRTFTVTMEGLRRDDADVWCGIERRGPDLTQVKVI
VDPEGAASTTASSPTNSNMAVFIGSHKRNHYMLLVFVKVPILLILVTAILWLKGSQRVPE
EPGEQPIYMNFSEPLTKDMAT
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-17

**Transmembrane domain:**

Amino acids 151-170

**N-glycosylation site:**

Amino acids 190-194

**Tyrosine kinase phosphorylation site:**

Amino acids 95-103

**N-myristoylation sites:**

Amino acids 66-72;125-131

**Prokaryotic membrane lipoprotein lipid attachment site:**

Amino acids 5-16



614/615

**FIGURE 609**

GATGGCGCAGCCACAGCTTCTGTGAGATTTCGATTTCTCCCCAGTTCCCCTGTGGGTCTGAGGG  
GACCAGAAGGGTGAGCTACGTTGGCTTTCTGGAAGGGGAGGCTATATGCGTCAATTCCCCAAA  
ACAAGTTTTGACATTTCCCCTGAAATGTCATTCTCTATCTATTCACTGCAAGTGCCTGCTGTT  
CCAGGCCTTACCTGCTGGGCACTAACGGCGGAGCCAGGATGGGGACAGAATAAAGGAGCCACG  
ACCTGTGCCACCAACTCGCACTCAGACTCTGAACTCAGACCTGAAATCTTCTCTTCACGGGAG  
GCTTGGCAGTTTTTCTTACTCCTGTGGTCTCCAGATTTAGGCCTAAGATGAAAGCCTCTAGT  
CTTGCCCTTCAGCCTTCTCTCTGCTGCGTTTTATCTCCTATGGACTCCTTCCACTGGACTGAAG  
ACACTCAATTTGGGAAGCTGTGTGATCGCCACAAACCTTCAGGAAATACGAAATGGATTTTCT  
GAGATACGGGGCAGTGTGCAAGCCAAAGATGGAAACATTGACATCAGAATCTTAAGGAGGACT  
GAGTCTTTGCAAGACACAAAGCCTGCGAATCGATGCTGCCTCCTGCGCCATTTGCTAAGACTC  
TATCTGGACAGGGTATTTAAAACTACCAGACCCCTGACCATTATACTCTCCGGAAGATCAGC  
AGCCTCGCCAATTCCTTTCTTACCATCAAGAAGGACCTCCGGCTCTCTCATGCCACATGACA  
TGCCATTGTGGGGAGGAAGCAATGAAGAAATAAGCCAGATTCTGAGTCACTTTGAAAAGCTG  
GAACCTCAGGCAGCAGTTGTGAAGGCTTTGGGGGAAGTAGACATTCTTCTGCAATGGATGGAG  
GAGACAGAAATAGGAGGAAAGTGATGCTGCTGCTAAGAATATTTCGAGGTCAAGAGCTCCAGTCT  
TCAATACCTGCAGAGGAGGCATGACCCCAAACCATCTCTTTACTGTACTAGTCTTGTGCT  
GGTCACAGTGTATCTTATTTATGCATTACTTGCTTCCTTGCATGATTGTCTTTATGCATCCCC  
AATCTTAATTGAGACCATACTTGTATAAGATTTTTGTAATATCTTTCTGCTATTGGATATATT  
TATTAGTTAATATATTTATTTATTTTTTGCTATTTAATGTATTTATTTTTTTTACTTGGACATG  
AACTTTAAAAAAATTACAGATTATATTTATAACCTGACTAGAGCAGGTGATGTATTTTTAT  
ACAGTAAAAAAAAAAAAACCTTGTAATTCTAGAAGAGTGGCTAGGGGGGTTATTCAATTTGTAT  
TCAACTAAGGACATATTTACTCATGCTGATGCTCTGTGAGATATTTGAAATTGAACCAATGAC  
TACTTAGGATGGGTGTGGAATAAGTTTTGATGTGGAATTGCACATCTACCTTACAATTACTG  
ACCATCCCCAGTAGACTCCCCAGTCCCATAATTGTGTATCTTCCAGCCAGGAATCCTACACGG  
CCAGCATGTATTTCTACAAATAAAGTTTTCTTGCATACCAAAAAAAAAAAAAAAAAAAAA

615/615

**FIGURE 610**

```
></usr/seqdb2/sst/DNA/Dnaseqs.min/ss.DNA83500
><subunit 1 of 1, 261 aa, 1 stop
><MW: 29667, pI: 8.76, NX(S/T): 0
MRQFPKTSFDISPEMSFSIYSLQVPAVPGLTCWALTAEPGWGQNKGATTCATNSHSDSEL
RPEIFSSREAWQFFLLLWSPDFRPKMKASSLAFLLSAAFYLLWTPSTGLKTLNLGSCVI
ATNLQEIRNGFSEIRGSVQAKDGNIDIRILRRTESLQDTKPANRCCLLRHLLRLYLDRVF
KNYQTPDHYTLRKISSLANSEFLTIKKDLRLSHAMTCHCGEAMKKYSQILSHFEKLEPQ
AAVVKALGELDILLQWMEETE
```

**Important features of the protein:****Signal peptide:**

Amino acids 1-42

**cAMP- and cGMP-dependent protein kinase phosphorylation sites:**

Amino acids 192-196;225-229

**N-myristoylation sites:**

Amino acids 42-48;46-52;136-142